

# The Hub in the Wheel ?

*Standardised tests as the pivot in the marketisation of education*

**Solveig Maria Lehoczky Gulling**



Master of Philosophy in  
Comparative and International Education

**UNIVERSITETET I OSLO**

Faculty of Education,  
Institute of Educational Research

Spring 2009



## Abstract

This study is about the neoliberalisation process within education, the process that also is called the *marketisation of education*. I have seen the process from a historical, political, economic, ideological and theoretical perspective, and used the PISA research to show the role of standardised tests in the marketisation process. I have chosen the metaphor *The Wheel* to illustrate this process; thus I have experienced the present development within the education sector as evolving, from a tiny movement in the beginning until the present where “everybody” moves in the same direction at high speed. *The hub* is the metaphor used for standardised tests. My research question asks whether standardised tests might be seen as the pivot in the marketisation of education, if they might act as *the Hub in the Wheel*.

To answer this question, I have conducted a critical discourse analysis of purposive picked documents. Own experiences, information from the media, at conferences, seminars and meetings, among people active in social movements and teachers working in the classroom has guided me in the search of suitable documents.

The study shows a number of parallels between the development within the economic and education sector. My findings indicate that standardised tests are important to the neoliberalisation process within the education sector. To make substantial conclusions, however, becomes more difficult the more one learns. A vehicle has more than one wheel, however not all are connected to the drive or to the breaks. I hope this study, this bricolage, shows that standardised tests are connected to the drive, and that to question them might lead to the release of some breaks.

## Acknowledgment

Writing a thesis like this require a lot more than the authors work. It is a process involving lots of people. First of all I will thank my friend, professor and supervisor, Birgit Brock-Utne, for encouraging me to apply for the program, for inspiring, knowledgeable and suggestive lectures and valuable tutoring undertaken in Oslo, Charlston and Cape Town. You have the ability to empower people, and I am grateful to be one of the many who have been able to benefit from this. You have always had faith in me; I hope this thesis will not let you down.

Then I will thank all the lectures of the CIE cohort of 2007-2009, and give a special thank to Halla Björk Holmarsdottir who invited me into what for me was the totally alien field of methodology. Moreover a special thanks to the PhD candidates Greta Björk Gudmundsdottir for always being helpful and Torill Agot Halvorsen for helping me to make the sketch showed in appendix J.

I will also thank my fellow students who had to bear with a classmate who of age could have been their mother. You have all contributed to two special years in my life. Two of you became closer. Thanks for all the interesting discussions, your company at all the exiting conferences and for your friendship, Zehlia. Thanks for your help to solve the mysteries of statistics and our conversations, my “daughter” from China, Lingyuan.

Two years in the CIE programme have been challenging, rewarding and fun. Due to Birgit, I got the possibility to attend NETREED 2007, IMPLAN 2008 and the LOITASA workshop 2009. Due to a grant given to me from Union of Education Norway I went to CIES 2008, CESE 2008 and CIES 2009.

I want to give a special thank to Union of Education Norway for the grant consisting of 30.000 NOK, and also for kind assistance when I needed it. A special thank goes to Anne Kathrine Blyverket who is a trade union official at the central office of Union of Education Norway. You have always been so helpful. Moreover, I will thank trade union representatives at the Oslo branch of Union of Education Norway for their support and understanding when the work with this thesis went at the price of my duty as a trade union representative.

---

This thesis is also indebted to all the people who have given me formal and informal inputs, and to the scholars from all over who have sent me papers and helped me to find literature according to my research question.

At last I will thank my knowledgeable, wise and helpful partner, Asbjørn. Your support has been of great importance.

# Table of contents

ABSTRACT.....	I
ACKNOWLEDGMENT.....	II
TABLE OF CONTENTS.....	IV
LIST OF TABLES .....	VIII
LIST OF ABBREVIATIONS.....	IX
<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 BACKGROUND .....	1
1.2 PROBLEMS AND STATEMENT .....	2
1.3 SIGNIFICANCE OF THE STUDY .....	4
1.4 SOME CONCEPTS.....	5
1.4.1 <i>Standardised testing</i> .....	5
1.4.2 <i>The marketisation of education</i> .....	7
1.4.3 <i>Some dichotomies</i> .....	7
1.5 LIMITATIONS .....	8
1.6 STRUCTURE OF THE STUDY .....	9
<b>2. THE HISTORY OF STANDARDISED TESTING .....</b>	<b>10</b>
2.1 THE OLD DAYS.....	10
2.2 FROM ELITE TO MASS SECONDARY EDUCATION .....	11
2.2.1 <i>International Examinations Inquiry (IEI)</i> .....	11
2.2.2 <i>The post war period</i> .....	13
2.2.3 <i>From the 1980s</i> .....	15
2.3 THE PRESENT SITUATION .....	16
2.3.1 <i>Why using PISA as an example?</i> .....	18
2.3.2 <i>The arguments for PISA</i> .....	19
2.3.3 <i>Some technical aspects of PISA</i> .....	21
2.4 SUMMING UP .....	23

---

<b>3.</b>	<b>THE DIALECTICS WITHIN AND THE CONTRADICTIONS BETWEEN.....</b>	<b>24</b>
3.1	IDENTIFICATION OF THE GUIDING THEORIES.....	24
3.1.1	<i>Marxism vs. Structural Functionalism</i> .....	25
3.1.2	<i>Dialectics</i> .....	27
3.1.3	<i>Critical theory</i> .....	28
3.1.4	<i>The production of knowledge</i> .....	29
3.1.5	<i>What kind of qualifications are “needed”?</i> .....	30
3.2	IDENTIFICATION OF DRIVING FORCES.....	31
3.2.1	<i>Globalisation</i> .....	32
3.2.2	<i>Neoliberalism</i> .....	33
3.2.3	<i>Neoconservatism</i> .....	37
3.3	THE EFFECTS OF THE NEOLIBERALISATION PROCESS.....	38
3.3.1	<i>The global effects</i> .....	38
3.3.2	<i>The effects on public services</i> .....	41
3.4	SUMMING UP.....	43
<b>4.</b>	<b>A BRICOLEUR.....</b>	<b>44</b>
4.1	A QUALITATIVE RESEARCH STRATEGY.....	44
4.1.1	<i>Research type</i> .....	47
4.1.2	<i>Data collection</i> .....	49
4.1.3	<i>Analysis</i> .....	52
4.1.4	<i>Challenges</i> .....	54
4.1.5	<i>Validity and reliability or trustworthiness?</i> .....	54
4.2	SUMMING UP.....	57
<b>5.</b>	<b>THE “WHEEL”.....</b>	<b>58</b>
5.1	LEFT VS. RIGHT.....	58
5.1.1	<i>The period of embedded liberalism</i> .....	59
5.1.2	<i>The period of neoliberalism</i> .....	60
5.1.3	<i>No differences between the left and right?</i> .....	61
5.2	TO CHANGE THE WAY PEOPLE THINK.....	62
5.2.1	<i>Think tanks</i> .....	63

---

5.2.2	<i>The use of test results</i> .....	64
5.2.3	<i>The use of test results in Norway</i> .....	65
5.2.4	<i>From Homo Politicus to Homo Economicus</i> .....	68
5.3	TO CHANGE WHAT PEOPLE DO .....	70
5.3.1	<i>Vital organisations</i> .....	71
5.3.2	<i>What kind of qualifications should be “produced”?</i> .....	77
5.3.3	<i>Business models as seen in education</i> .....	80
5.4	WHAT KIND OF WHEEL, WHAT KIND OF HUB?.....	84
<b>6.</b>	<b>THE “HUB”</b> .....	<b>86</b>
6.1	SETTING STANDARDS.....	86
6.1.1	<i>Technical challenges</i> .....	87
6.1.2	<i>A question about will and power</i> .....	88
6.2	WHOSE KNOWLEDGE COUNTS?.....	90
6.2.1	<i>The test items of PISA</i> .....	91
6.3	WHAT KIND OF KNOWLEDGE COUNTS? .....	98
6.3.1	<i>What kind of knowledge and skills counts for the students?</i> .....	98
6.3.2	<i>What kind of knowledge is produced by PISA?</i> .....	100
6.3.3	<i>A methodological straightjacket?</i> .....	107
6.4	WHY?.....	109
<b>7.</b>	<b>THE BRICOLAGE</b> .....	<b>111</b>
7.1.1	<i>The contradictions between</i> .....	111
7.1.2	<i>The connections within</i> .....	113
7.1.3	<i>A Hub in the wheel</i> .....	115
<b>8.</b>	<b>REFERENCES</b> .....	<b>117</b>
	<b>APPENDIX A</b> .....	<b>128</b>
	<b>APPENDIX B</b> .....	<b>129</b>
	<b>APPENDIX C</b> .....	<b>131</b>
	<b>APPENDIX D</b> .....	<b>132</b>



---

<b>APPENDIX E .....</b>	<b>133</b>
<b>APPENDIX F .....</b>	<b>134</b>
<b>APPENDIX G.....</b>	<b>135</b>
<b>APPENDIX H.....</b>	<b>136</b>
<b>APPENDIX I .....</b>	<b>137</b>
<b>APPENDIX J.....</b>	<b>138</b>

## List of tables

<b>Table 2.1:</b> An overview of important educational assessments in use.....	17
<b>Table 3.1:</b> Modernist theories.....	26
<b>Table 3.2:</b> Postmodernist and poststructuralist theories.....	26
<b>Table 3.3:</b> Qualification framework.....	31
<b>Table 5.1:</b> The Uruguay agreement as seen in education.....	72
<b>Table 7.1:</b> Contradictory and dialectical concepts.....	112

## List of abbreviations

ACER – Australian Council for Educational Research

CDA – Critical Discourse Analysis

CERI – Centre for Educational Research and Innovation

DA – Discourse Analysis

EI – Education International

EFA – Education for All

EKVA – Enhet for Kvantitative Utdanningsanalyser [Unit for Quantitative Analysis of Education]

ESAP – Economic Structural Adjustment Programmes

EU - European Union

GATS - General Agreement on Trade in Services

GNP – Gross National Product

HDI – Human Developing Index

ISCED – International Standard Classification of Education

IEA – International Association for the Evaluation of Educational Achievement

IEI – International Examination Inquiry

ILS – Institutt for Lærerutdanning og Skoleutvikling [Department of Teacher Education and School Development]

IMF – International Monetary Fund

NAEP – National Assessment of Educational Progress

NGO – Non Governmental Organisation

NOK – Norske Kroner [Norwegian Kroner]

NPM – New Public Management

OECD – Organisation of Economic Cooperation and Development

PIRLS – Progress in Reading Literacy Studies

PISA – Programme of International Student Assessment

RCT – Randomized Controlled Trials

ROR – Rate of Return

SIMS – Second International Mathematics Study

SISS – Second International Science Study

SSA – South Saharan Africa

SSØ – Senter for Statlig Økonomistyring [Government Agency for Financial Management]

TIMSS – Trends in Mathematics and Science

TINA – There Is No Alternative

UiO – University of Oslo

UK – United Kingdom

UN – United Nations

UNESCO – United Nations Educational, Scientific and Cultural Organization

UNDP – United Nations Developing Programme

US – The United States of America

USSR – The Soviet Union

WTO - World Trade Organisation

WWI – World War I/the First World War

WWII – World War II/the Second World War

# 1. Introduction

## 1.1 Background

As a teacher from the late 1970s, I have been part of the development within the education sector in Norway. I have worked within the frame of four different curriculum planes, under the guidance of a number of different White Papers on Education written by different Ministers from different political parties and under different political County and Municipality regimes. During those years, I have experienced a rather profound shift that has influenced all levels of the education system; the pupils; the teachers; the administration and the organization as a whole. Taken the time span and the general development of our society into consideration, a development within the educational system is not surprising. At one point, however, I do not remember exactly when, I started to feel that the development was going in reverse.

Most of the time I have been working in Oslo, where steps were taken to “modernize” the steering structure in the 1990s; the economic allotment system, the account system and the planning and the evaluation system. Earlier the headmaster was the first among equals, now she is the extended arm of the central administration. Earlier the economy was predictable; the allotment was given according to the school year and class size. Now the economic situation is unpredictable; the money follows the pupils, the pupils are counted twice a year and the allotment is calculated according to those figures. Earlier the students were assigned to the neighboring school, now the pupils have a free choice. Earlier there were no central standardised tests, now there are several; Oslo-, national- and international tests, and the results of the tests are published. Earlier there were central agreements defining maximum number of students in class, now there are no such limits. Earlier the teachers’ salaries were decided by central agreements; now more and more money is given to local salary negotiations. Earlier the working hours of teachers were settled by central negotiations between the authorities and the unions. In the future we are facing a situation where also the working hours might become negotiated at local level. Earlier a teacher devoted most of the time to the students, now time is used for bureaucratic work like filling in forms and evaluations of all sorts. Moreover precious time is spent in endless discussions defined by

the central administration. Earlier our strategic plans and goals were defined according to the need of the school and formulated in words, now they are defined by the central administration and formulated in numbers<sup>1</sup>.

In the beginning, I thought that the changes were due to a misunderstanding, that the politicians took all the “wrong” decisions because they did not know. If only the teachers, the unions and the educational researchers were able to come in a dialogue with those in power, we could give them input to guide them. After a while, however, I started to realize that all the small steps were guided by ideological, political, epistemological, ontological and methodological positions, and that the steps were taken deliberately. In retrospective we can see that what started with small, at the time they were implemented rather innocent steps, all of a sudden stood out as a systemic shift. Seen from a teacher and a unionist perspective, those changes have lead to growing frustrations. Seen from the central school administration and the politicians ruling in the city, the system is a success story that they constantly try to export to other counties and municipalities in Norway.

During my process of awakening, I have asked lots of questions, many to which there are no obvious answers. Put in another way; the answers depend on what kind of glasses one wears. Nevertheless, I am still asking, and one of my huge questions is connected to the driving forces of the present development. Another is whether it is possible to identify some concrete measures that are more important for the development to continue than others.

## 1.2 Problems and statement

### *Main objectives*

There are heaps of international and national literature about the present development within education. I have, however, not found anyone taken the same point of departure as I have done by my main research question: Do standardised tests act as the pivot in the marketisation of education – can they be seen as *the hub in the wheel*?

---

<sup>1</sup> The information given here can be seen as a compressed sum up of countless informal and formal discussions between teachers at school level and representatives from the Oslo branch of Union of Education Norway (In Norwegian: Utdanningsforbundet Oslo).

---

The need to learn more about standardised tests and to put them into a broader context emerged while working with my term paper the spring of 2008. Questions like these emerged: Within what kind of epistemological, ontological, ideological and political culture are standardised tests born? What kind of role do they have in the development of the education sector? How does the use of them influence different aspect of schooling? Is it possible to see the rise in international and national tests as a conscious step towards more market oriented educational systems all over the world?

To solve my task I have chosen a broad, theoretical entrance to the issue, building on international and Norwegian literature and examples. To ease my work, I have split my main research question into the following questions.

*Research questions:*

1. What is the history behind the increasing use of standardised tests internationally?
  - a) Where and why did they start to use them?
  - b) What kind of ideology and political forces promote them?
  - c) How and why did they disperse?
2. What is the history behind the test culture in Norway?
  - a) When did we start to join the international test programs?
  - b) Why did we start?
  - c) What have been the effects?
  - d) Who promotes the development of a test culture and who opposes it?
3. What are the effects of standardised tests on
  - a) what kind of knowledge that is produced
  - b) whose knowledge counts
  - c) what type of qualification that counts
  - d) profit making?

### 1.3 Significance of the study

There has been a growing dissatisfaction with the present development of the education sector in Norway. The teacher unions have not been able to fight the development. It seems that those promoting it have been better prepared and organized than those opposing it. We have been protesting, may be occasionally managed to postpone, but not managed to stop the development. The same has happened all over the world. It has been like a wave, difficult to grab and to withstand. Some call it globalisation implying *There Is No Alternative*<sup>2</sup>.

However what have been made by humans may be changed by humans. The first step is to name the world (Freire 1993) to identify the different parts of the wave and the connection between them. As I have come to see it, the test culture might be an essential part. Therefore, to dive into essential documents about the ideas behind, the implementation and the results of the testing culture may be one of more starting points in the struggle against the marketisation of education. The former regional coordinator of Africa within *Educational International* (EI), Tom Bediako<sup>3</sup> in his last speech in Norway before he retired, urged us to start to fight the market liberal forces in our own country. “If *you* loose this fight”, he said, “Africa is lost for ever”. As I see it, to be able to fight nationally and locally, it is important to learn from experiences elsewhere, to take the dialectic between the local and the global seriously.

Giroux (2003), Gorostiaga (1993) and Ki-Zerbo (1994) underline the social responsibility of the scholars in a diverse world. Brock-Utne (2006) was in a position to write a book in order to contribute to the uphill struggle against the World Bank. As a unionist I wanted to use this opportunity to gather knowledge of importance in order to contribute to our local and national uphill struggle against the present development within the education sector.

---

<sup>2</sup> The well known saying of Margaret Thatcher: *There Is No Alternative* (TINA).

<sup>3</sup> From the conference *Strategies in international solidarity work seen from the south*, Clarion hotel Oslo airport, June 2002, conducted by Union of Education Norway.



---

## 1.4 Some concepts

In this section I will introduce some concepts which are important to explain the title of this thesis. They are also important for the discussions in chapters to come.

### 1.4.1 Standardised testing

In the title I use the term “standardised tests”. To clarify what I mean, I will define some essential concepts.

#### *Standards*

In their article, Husén & Tuijnman (1994: 2) define standards with the following words of Livingston (1985): “A *standard* refers to the degree of excellence required for particular purposes, a measure of what is adequate, a socially and practically desired level of performance”. Husén & Tuijnman identify three types of educational standards; *opportunity to learn standards*, *content standards*, and *performance standards*.

#### *Assessments*

“Assessment refers to the techniques used in collecting information about educational outcomes either subjectively by using experienced judgements or by means of standardized, objective tests”<sup>4</sup> (Husén & Tuijnman 1994: 3). Most often assessments evaluate learners’ outcomes, personality or ability. Cummings (2003) presents four main types that are used in education:

- Assessments testing knowledge. They are closely connected to the content standard, testing what the learners remember from their schooling.
- Assessments testing skills. They are closely connected to performance standards and test the ability to use knowledge.
- Assessments testing values, attitudes or convictions.
- Assessments testing aptitudes.

---

<sup>4</sup> Whether standardised tests are objective or test what they are supposed to do, will be dealt with later in this text.

### *Evaluation, monitoring and accountability*

Students' assessments can be used as instruments in both evaluation and monitoring processes. Not only learners are being evaluated and monitored, however. Within the education system we can find evaluation and monitoring activities at all levels; student, teacher, headmaster, school, municipality, county, country and cross country level. Whereas evaluation is defined as "the systematic collection and interpretation of evidence, leading, as a part of the process, to a judgment of value with a view to action" (Beeby 1977 in Husén & Tuijnman 1994: 3), monitoring "refers to systematic and regular procedures for the collection of data about important aspects of education at national, regional or local levels" (Husén & Tuijnman 1994). In other words, evaluation uses the data collected from assessments to make a value judgment of an individual, while "[m]onitoring involves the systematic collection of evidence about the context, inputs, processes and outcomes of an education system" (ibid.: 3-4). Monitoring is by this definition a wider concept than evaluation, taking more than performances into consideration.

*Accountability* in the education sector means that stakeholders, from parents to governments should be able to check whether a school or a school system functions. What schools produce should be accounted for. Accountability is supposed to "enriching public discussion by reporting on the overall status and strengths and weaknesses in education, thus encouraging the setting of education goals and performance standards" (Husén & Tuijnman 1994: 16).

### *Standardised tests*

Evaluations can be *formative* or *summative*. Cummings (2003: 194) defines formative evaluation to be "continuously conducted during the students' learning activities" and summative evaluation to be "assessment conducted at the end of a lengthy period of study, such as at the end of a semester, of a year, or of a major segment of schooling hierarchy". While formative evaluation can be based on all kinds of assessments; those made by the teacher or standardised tests, summative evaluation is often dependent on pre-produced tests, or standardised tests. If they tests knowledge, they are made in association with a curricular plane or syllabus. If the aim is to test skills, those making the tests have to consider what a learner at that particular level should be able to perform. There are also lots of standardised tests testing values, attitudes and aptitudes. As with the other tests, they are all made to meet

---

a standard; a certain level or conception of what should be the way of acting, behaving or thinking.

### 1.4.2 The marketisation of education

When I use the term *the marketisation of education* in the heading of this thesis, I refer to the development within the education sector that started a quarter of a century ago, and that has been implemented with different strength in different countries during the same period. This is a contested policy; the content of my thesis lies in the tension between the rhetoric of this policy and the effects of it.

### 1.4.3 Some dichotomies

The following dichotomies will be used to ease my analysis later in this study. They will consequent be written in italic.

#### *The north and the south*

Yearly the *United Nations Developing Programme* (UNDP) launches their *Human Developing Report* where the *Human Developing Index* (HDI) of all *United Nations* (UN) member states are divided into three categories; “low”, “middle” and “high” human development (UNDP 1990). When I use the concepts *north* and *south* in this paper, I am not referring to geographic areas, but to areas defined according to the categories of human development. Accordingly *north* refers to the “developed” countries; the countries in the category of “high” and upper part of “median” human development. *South* refers to the “developing” countries; the countries in the category of “low” or lower part of “median” human development.

#### *The uppers and the lowers*

Chambers (1997) divides the people in the world in *uppers* and *lowers*, a differentiation that distinguishes between those holding a high or low status due to education, training, induction, competence, gender, influence, wealth and location. This covers the differences *within* the *north* and the south. While the *uppers* represent the majority of the *north* and the minority of the *south*, the *lowers* represent the minority of the *north* and the majority of the

*south*. Hence; we can find a *south* in the *north* and a *north* in the *south*, the class society is found everywhere.

#### *The left and the right*

The *left* represents those holding political radical positions; socialists, communists. The *right* are political conservatives; capitalists, neoliberals.

#### *The west and the east*

Those concepts are connected to the *post war* period, where the *west* represents the western part of the world and also the *right*, and the *east* represents the countries claiming to be socialists or communists, and hence the *left*.

#### *The insiders and the outsiders*

In this thesis I have chosen the terminology the *insiders* and the *outsiders*. The *insiders* are engaged in the *Programme of International Student Assessment (PISA)*, the *outsiders* are critical to the study in one way or another

## 1.5 Limitations

My main limitation has been the time frame seen in connection to the complexity of my research question and the amount of literature available on the issue. The task has been growing and growing parallel to my reading. Each time I felt I had come to data saturation, something new appeared. I blame this experience on the lack of research experience. I must admit that if a new crossroad will appear, I will narrow down my research question.

Another limitation has been that my main research question has been too abstract, and hence difficult to answer. Accordingly, the lesson I have learnt is to make the research question more concrete. However, everything becomes clear in the light of belated wisdom.

Another limitation might be my situation; an experienced teacher and unionist researching on an issue highly connected to her profession and position. Some will argue this is a bias, and hence question the trustworthiness of the research. Nevertheless, a pretext to discuss a task most often is not regarded as a drawback in the real life; rather the opposite. Kvale (1996:182) values experience as well:

---

The importance of background knowledge for observations is evident in a variety of areas. When analyzing interviews with chess players, the researcher's knowledge of chess at a higher level than that of the interviewees is a precondition for seeing the solutions they did not see.

We all enter into a task carrying a luggage; to call this a bias is a bias in itself. To claim to be objective is a bias, since it is to shut the eyes for one's own subjectivity. What is important is to be aware of and open about one's position, to be self reflexive.

## 1.6 Structure of the study

This study consists of *seven* chapters. *Chapter I* set the stage. It presents the background and the main objectives of the study. Moreover, it discusses its significance and limitations, and defines some concepts and dichotomies. Some of the concepts facilitate the comprehension of the heading and some prepare for later discussions. The dichotomies will be used to ease the analysis to come.

*Chapter II* gives an historical overview of the development of standardised tests, from the Sui Dynasty until today. This chapter also covers PISA technicalities and explains the choice to use this assessment as an example in the discussion of standardised tests.

*Chapter III* presents the theoretical framework and the economic and political context that will guide my analysis in chapter V and VI.

In *chapter IV* I present myself as a bricoleur, the quilt-maker, that intend to gather bits and pieces and put them together into a new whole. This is the methodology chapter, where the choice of research strategy, research type, data collection and analysis, as well as challenges will be presented.

*Chapter V, The Wheel* and *chapter VI, The Hub*, are the main chapters. There data is presented and analysed. Chapter V aims to look into the process of marketisation of education from a historical, political, ideological and economic perspective. In chapter VI I will use PISA as an example and look into theoretical perspective of standardised tests, the epistemological, ontological and methodological perspectives.

In chapter VII my aim is to conclude, to finish my *bricolage*.

## 2. The history of standardised testing

My main objective in this study is to find out if standardised testing can be seen as a pivot in the marketisation of education. However, before I can approach this difficult question, I have to understand why a culture where the need to evaluate and monitor input, output and processes through the use of standardised tests has been developed. A starting point is to look into the history. Hence, in this chapter I will use existing literature in order to make an historical overview of the development of standardised testing. I will pose some questions, but leave the discussion to later chapters.

### 2.1 The old days

The tradition to test people has a long history. According to Cummings (2003) it is believed that the written form of testing, the examination, originated in the Sui Dynasty in China (AD 589 – 618) and that the tradition was brought to Europe by the Jesuits in the early 1700s. Before that the testing was an oral exercise in Europe, built upon the rhetorical Judaeo-Hebraic tradition that had been further developed by Platon.

The reason for the written examination was to select people for special tasks. During the Sui Dynasty it was the wish to select officials to the imperial bureaucracy from outside the aristocracy they had inherited from the former dynasty. In Prussia the reason to use examinations was to pick out governmental officials, and when they established the *Abitur* in 1788, it was to select people for the universities. Later on examinations also were introduced at lower levels. Although the different countries developed their own systems, Prussia started an examination tradition that was to become common during the nineteenth century in the western hemisphere. Also Japan in the mid-nineteenth century as well as the Soviet Union (USSR) in the beginning of the twentieth century followed, all according to Cummings (2003).

## 2.2 From elite to mass secondary education

Except for the USSR and China, that did not join the industrial revolution in full before after their own revolutions in respectively 1917 and 1949, all countries mentioned in 2.1 were part of the industrial revolution in the nineteenth century. When their industry matured the demand for skilled workers rose as did the demand for civil servants alongside the growing governmental bureaucracy. These are both occurrences that increased the need to expand secondary education as well as to select those entering this level. Around 1900, the *United States of America* (US) were in such a position. According to Cummings (2003) a rapid growing industry and a need to select thousands of workers with backgrounds from different cultures and countries was the starting point of psychological testing, which is the use of assessments testing abilities, personalities and intelligence. After being used in the recruitment of military service personnel during the First World War (WWI) testing soon after entered the school system in the US.

This development of psychological testing with the aim to select, led to the development of psychometric instruments that enable us to combine and compare input and output of schools. The first time they used evaluation to compare the outcome of different schools through the testing of cognitive performances as well as different affective characteristics of the students, was in the US in the 1930s (Husén & Tuijman 1994). However, it took years before the psychometric tests became fully developed as instruments to be used in monitoring on a big scale, and years before such kind of monitoring became common in Europe and in the rest of the world. How did the ideas spread from the US?

### 2.2.1 International Examinations Inquiry (IEI)

IEI was a scientific research project in the 1930s funded by Carnegie Corporation. Originally it comprised of researchers from England, France, Germany, Scotland, Switzerland and the US. In the late 1930s Finland, Norway and Sweden were added. Those countries and the scholars to participate were chosen because of previous educational work and personal relationship (Lawn 2008 a). They represented three different disciplinary interests; “founders and popularisers of the study of intelligence (....), the first generation of comparativists in education (...) and key members of a world progressive education movement” (Lawn 2008a: 11). The point of departure for the research was the expansion of secondary education, which

was on its way of moving from elite to a mass education system. Scholars in many European countries questioned the old examination system and were looking for new selecting mechanisms. The funding by Carnegie provided opportunities to dive into this issue, to meet and share ideas and findings and to be exposed to the new American solutions. What Carnegie funded was research projects in the participating countries and conferences in 1931, 1936 and 1938 where the research findings were presented and discussed.

The book, *An Atlantic Crossing?* (Lawn 2008b) is written by present authors from all the countries that participated in IEI, using national files from the days of the *Inquiry* as the base of their articles. The book questions whether the new psychometric testing ideas and instruments were imported from the US to Europe or if the appearance of this type of testing simply was a result of an independent European development. According to Lawn (2008a: 8) historians in former colonies “are more aware of the travel of technologies or systems of pedagogy, and of the centre-peripheral relations” than historians dealing with educational history in the western part of the world. This explanation, but also national differences and different understanding of the issue leave us without a plain answer to the question above. Lawn (2008 a) is quite clear, however, he argues that a stream of influence has come across the Atlantic. First of all, using English as a working language reflects “the beginning of an internationalism that reflected a new axis, America” (ibid.: 12). At that time, English was not the common academic language in mainland Europe. Secondly this new axis had “a clear sense of mission to modernise examining practices in other countries” (ibid.: 20). Thirdly IEI was a medium for the empirical – experimental research movement in Europe (ibid.). A part of the technology developed through the work of IEI was what Lawn (ibid: 14) calls “the new common language of education research, a kind of scientific Esperanto”. He refers to concepts like tests, surveys, statistics, scientific objectivity, reliability, universal standards etc. IEI was in other word a starting point for educational quantitative, statistical research:

It promised a usable truth through numerical analyses. Through the alchemy of their procedures, with a language of reliability and emerging discourse of objective and universal standards, they were beginning to establish their authority over the older cultural connoisseurship and qualitative judgments of the earlier internationalists (ibid: 15).



## 2.2.2 The post war period

A seed to a new understanding of evaluation in our part of the world was sown. However, WWII delayed the nurturing, but in the early post war period the seeds started to sprout in some of the countries that participated in IEI: In England the work of the *English Committee* of IEI “became the context for the post war development of the English education system” (Lawn 2008c: 58). In Scotland the *Scottish Council for Research in Education* “formed itself within an American school of thought on educational research” (Lawn, Deary & Bartholomew 2008: 119) and had huge impact in the post war period. In Sweden “the post-war education reforms were influenced by the work of (...) members of the Swedish research project” (Lawn 2008a: 27) and was the only country in Europe that had done comparative evaluation studies of schools before the mid 1950s (Husén & Tuijman 1994: 4). In Germany Lawn (ibid.: 27) mentions that Hylla, one of the German researchers “in the early 1950s (...) translates American key test terms into a German glossary”, however Waldow (2008: 61) argues that “[c]onceiving of education as *Bildung*<sup>5</sup>, (...) potentially stood in conflict to applying psychometric testing methods”.

In other countries that participated in IEI the seeds were put to rest, as in the case of Switzerland and France. While the important participant professor Pierre Bovet from Switzerland was highly ambivalent to the issue of testing (Hofsetter & Schneuwly 2008), the French Committee “failed in sustaining the international efforts to improve European countries’ way of examining” as Zarrouati (2008: 99) puts it, and “its work sank in the deep waters of indifference and oblivion” during World War II (WWII). In the case of Finland the work of IEI also had limited impact (Vuorio-Lehti & Jauhiainen 2008: 152) and in Norway the “use of testing as a functional alternative of supplement to a professional assessment and examination was a contested position” (Jarning & Aas 2008: 198). In the early 1960s the further development of a test culture was rejected in the Norwegian Parliament, as noted by the same authors. In the immediate post war period those countries did not emphasize on

---

<sup>5</sup> *Bildung* is a German word; in Norwegian we use the word *dannelse*. There is no equivalent word in the English vocabulary. To give a pupil or a student *Bildung* is to give her an education that has a value in itself, regardless of economic needs. This view on education might stand in opposition to the use of standardised tests.

data collection connected to education, nor did other countries. *If* data was collected, it was to evaluate the input (Husén & Tuijman 1994).

In the US, however, the seeds were blooming, at least for a period. After the war *Educational Testing Services* developed *Scholastic Aptitude Tests* (SAT) to facilitate the college application process. The explanation given for the need for such tests was that during the war a whole generation of young American boys got their education interrupted and did not have common knowledge standards to assess while coming back (Cummings 2003). However, this was the case in many other countries as well. A following question is thus why it did not lead to the same development in other countries?

The Sputnik shock in 1957 was followed by a demand to improve the American education system. The *National Defence Education Act* of 1959 was *the* answer and “marked the entry into a new era in framing education policy at the national level in the United States, and later in Europe” (Husén & Tuijman 1994: 5). Parallel to this, the demand for more knowledge about the educational system rose. Now new techniques, methods and instruments that enabled the researchers to conduct sample surveys, to test groups, to draw inferences from the sampled population to the target population, to measure and analyse came into use. The instruments had been developed by psychometricians in the US and *United Kingdom* (UK) and social scientists in the US. Soon it was possible to evaluate input, process, output and progress, and in 1969 an assessment meant to produce a sort of educational Gross National Product, the *National Assessment of Educational Progress* (NAEP) came into use. Husén & Tuijman (1994: 6) argue that the combination of the human capital theory<sup>6</sup> and the evaluation techniques “made it possible in the early 1960s to embark on the development of what was to become an empirical approach in comparative education”.

According to Husén & Tuijman (1994) those following in the steps of the US during that time, were organisations like *United Nations Educational, Scientific and Cultural Organization* (UNESCO)<sup>7</sup> and *Organisation of Economic Cooperation and Development* (OECD). The UNESCO Institute in Hamburg played an instrumental role in promoting

---

<sup>6</sup> I will discuss the human capital theory in chapter III.

<sup>7</sup> Isaac Kandal, Professor of Education at Teachers College, Colombia University was a central actor in IEI. In 1944 he argued that the International Education Organisation, the precursor of UNESCO, should build on the same model as IEI. Kandal later became a major advisor in UNESCO (Husén & Tuijman 1994, Lawn 2008).

evaluation in Europe (ibid: 5). At a meeting in 1958<sup>8</sup> it was “proposed that one should try to conduct strict comparisons between national systems of education using empirical measures of resource inputs and student achievement” (ibid: 5). This proposal leads to the establishment of *International Association for the Evaluation of Educational Achievement* (IEA) in 1961. In 1973 OECD presented a framework of 46 indicators to guide governmental decisions, in the mid 1970s *International Standard Classification of Education* (ISCED) was implemented and in 1978 the US congress by law required periodic national assessments.<sup>9</sup>

Nevertheless, something happened in the early 1970s in the US and in the rest of the world. The blind faith in social engineering faded, and within the research area “controversy over systems theory and its linear models of reality in many circles stifled the belief in a fact-finding mission of the social sciences” (Husén & Tuijman 1994: 9). The indicators of OECD did not come into use, the pace of NAEP was slowing down and the rigidity and decontextualised information in ISCED was criticized: “Since a country’s education system is deeply rooted in its history and part and parcel of its socio-cultural matrix, a classification useful for interpreting the results is of limited use” (ibid.: 9). At the same page they state: “By the mid- 1970s, macro-level educational planning had become highly suspect; instead the capacity of education systems to resist structural and procedural reform was emphasized”, and the ambitious agenda of UNESCO and the OECD “for coordinated international data collection in education was abandoned”. What happened in the 1970s is another question to be dealt with in chapters to come.

### 2.2.3 From the 1980s

In 1983, the report *A Nation at Risk* and the result of the second international IEA study of mathematic was published in the US. In the report, the American school system was criticised and the IEA results showed that both the US and the west came out as losers

---

<sup>8</sup> A small curiosity: Prof. Edward Thorndike was a central participator in IEI (Lawn 2008). His son, Robert Thorndike was one of the participants at this meeting (Husén & Tuijman 1994), and is regarded as one of the fathers of IEA [http://www.iea.nl/brief\\_history\\_of\\_iea.html](http://www.iea.nl/brief_history_of_iea.html) (Accessed 06.01.09). The description of their relationship, see: <http://www.indiana.edu/~intell/rthorndike.shtml> (accessed 15.05.09).

<sup>9</sup> Another curiosity: Friederick Keppel was the one initiating IEI as the president of Carnegie Corporation (Lawn 2008). His son, Francis Keppel was the one that initiated this law (Husén & Tuijman 1994). The description of their relationship, see: <http://www.answers.com/topic/keppel-francis> (accessed: 01.11.08).

compared to Japan. As noted by (Husén & Tuijman 1994: 10), the way the report and the results were used, helped to change the focus of the 1960s and 1970s:

The *A Nation at Risk* report played a pivotal role in raising concern over the quality of education. In a more subtle way it also raised political awareness that the international comparative studies of student achievement could be used for the benchmarking of performance levels in different education systems.

In the US the *National Education Goals Panel* was appointed in 1990, with the purpose “[t]o hold the nation and the states accountable for their attainment” (ibid.:12) and the *National Council on Education Standards and Testing* (NCEST) was established in order to develop nationwide educational standards. “Ambitions of this kind have served to enhance the interest of several governments (...) to support international surveys of student achievements” at a time when a “new, system-wide framework influenced thinking about the nature of educational evaluation” (both quotations: ibid.: 12).

In 1984, the Ministers of Education of the OECD countries met. “The shifting of priorities from issues in managing the quantitative growth of the education system to cost-effectiveness and quality assurance was striking” (Husén & Tuijman 1994: 11). According to them, questions showing a milestone in the way to perceive education were asked: Do the ‘customers’ get ‘value for their money’? Is the public system efficient enough? A feasibility study done by OECD to investigate whether it is possible to “develop a limited set of *indicators* of education system”, led to the “decision to produce a limited set of international comparisons on key aspects of the education system” (ibid.: 13, italics is original).

## 2.3 The present situation

Husén & Tuijman (1994) argued in 1994 that the demand for information about the outcome of schooling and the political will to use the methods in order to feed the society with the information they demand, were in place. “[E]ducation policy has lost some of its preoccupation with localism and regionalism”, and “the field [of comparative education] is now approaching the time when hierarchical, multilevel information system in education can be established” (Husén & Tuijman 1994:13). Seen from a 2009 perspective, we know that they were right, “[t]he number of national and international assessments of learning outcomes has raised significantly” (UNESCO 2007: 26).

**Table 2.1. An overview of important educational assessments in use.**

Level	Where	Organised by	Name of the tests	Additional info.
<b>International assessments</b>	Worldwide	IEA	TIMMS	Appendix B
			PIRLS <sup>10</sup>	Appendix C
			MORE	Appendix A
		OECD	PISA	Appendix D
			MORE	
<b>Regional assessments</b>	“Anglophone” <sup>11</sup> countries in Africa	SACMEQ <sup>12</sup>	SACMEQ	
	“Francophone” countries in Africa	PASEC/CONFEMEN <sup>13</sup>	PASEC	
	Latin America	Coordinated by UNESCO	LIECE <sup>14</sup>	
<b>National assessments</b>	123 countries are listed in Global Monitoring Report 2008	Different	Different	Global Monitoring Report (2008: 209-220)
	Norway	National Quality Assessment System <sup>15</sup>	National Tests	Appendix E
			Mapping tests	

<sup>10</sup> *Progress in Reading Literacy Studies* (PIRLS).

<sup>11</sup> It is common to divide Africa into Anglophone, Francophone and Lusophone countries according to the language of the former colony powers. When I put “Anglophone” and “Francophone” countries in quotation marks it is to show that the terms are artificial, since less than 10% of the population speaks respectively English and French in those countries (Brock-Utne 2006).

<sup>12</sup> *Southern & Eastern Africa Consortium for Monitoring Educational Quality* (SACMEQ). SACMEQ cooperate with IEA and is involved in the *World Bank’s Global Partnership for the Assessment of Educational Achievement* : [http://gb.iiep-unesco.org/public/format\\_long\\_en.php?format=court&fiche=133](http://gb.iiep-unesco.org/public/format_long_en.php?format=court&fiche=133) (accessed 02.01.09).

<sup>13</sup> Programme d’analyse des systèmes éducatifs (PASEC ) and Conférence des Ministres de l’Education des pays africains et malgache d’expression française (CONFEMN): <http://www.confemen.org/> (accessed 02.01.09).

<sup>14</sup> *The Latin American Laboratory for Assessment of the Quality of Education* (LIECE): <http://llece.unesco.cl/ing/acerca/> (accessed 02.01.09).

<sup>15</sup> A concept used in Global Monitoring Report 2009: 219.

---

Local assessments	Oslo	Local school authorities	The Oslo tests	
-------------------	------	--------------------------	----------------	--

### 2.3.1 Why using PISA as an example?

“In summary, the reason why monitoring came about are closely linked to the perceived functions of monitoring national systems of education” Husén & Tuijnman (1994: 16) argue. Dolin (2007) states that international tests are the lever to the introduction of local, national and regional tests; that tests breed tests. Hopmann (2007) indicates that the introduction of the international tests together with the demand for accountability made way to national tests in many countries, and Uljiens (2007: 299) while writing about PISA asks: “What if the idea was rather to use international evaluation as a technique for homogenising the participating educational system and creating a competition-oriented mentality”?

The presumed important role of the international tests is the reason why I in the following will use one of these tests as an example when discussing standardised tests. Although it might have been better to exemplify through *Trends in Mathematics and Science* (TIMSS) because of its longer history<sup>16</sup>, I will use PISA. The reason is that the latter has become very influential (Langfeldt 2007, Sjøberg 2009).

This is especially truth in the case of Norway. The PISA 2000 results were for one thing used to explain the demand to establish national tests. At that time the establishing of national tests was controversial, however the *centre/right* government in power, managed to get support for the idea in the Parliament (Bergesen 2006)<sup>17</sup>. When the first *National Tests* were accomplished during the spring of 2004 (Kunnskapsdepartementet 2003)<sup>18</sup>, there were huge protests, and the year after they grew (Bergesen 2006). To make a long story short, due to the following occurrences the tests were stopped: An evaluation carried out by the *Norwegian*

---

<sup>16</sup> See appendix B.

<sup>17</sup> All translations from Norwegian and Danish in this thesis are done by the author.

<sup>18</sup> In English: Norwegian Ministry of Education and Research (official translation).

---

*Directorate for Education and Training*<sup>19</sup> that revealed substantial weaknesses with the tests (Bergesen 2006), massive protests by pupils supported by *The Pupils Organisation*<sup>20</sup> (ibid.), reactions among the teachers as shown in a survey initiated by the *Union of Education Norway* (TNC Gallup 2005), and the elections in 2005 when the *centre/right* government lost their power. After improvements and an agreement with the teacher union not to make the result public in order to compare the results between schools, they were reintroduced in 2007 by the new Government<sup>21</sup> (Utdanningsdirektoratet 2007). See appendix E for more information about the tests.

Parallel to the introduction of national tests, *Mapping tests* were introduced at the national level, and the *Oslo tests* in Oslo. There have not been the same reactions towards the Mapping tests, probably since they are not used as a means to external comparison, although teachers in Oslo are concerned about how they are used in internal comparison within schools and as governing tools for the Oslo school authorities<sup>22</sup>. The Oslo tests, however, have caused huge debates in the Oslo media and also within the teacher union because the results are being published.

The limitations of this thesis do not allow me to go into details about the Norwegian tests, although it could have been very interesting. They are just mentioned to illustrate the growing number of standardised tests at all levels.

### 2.3.2 The arguments for PISA

The PISA study was the direct result of the OECD meeting in 1984 and the following feasibility study (cf. 2.2.3). *Measuring Student Knowledge and Skills* (OECD 1999) describes the arguments for PISA, the design and assessment framework of the study. The document starts as following:

---

<sup>19</sup> In Norwegian: Utdanningsdirektoratet (official translation).

<sup>20</sup> In Norwegian: Elevorganisasjonen (my translation).

<sup>21</sup> This is a *centre/left* government consisting of a coalition of three political parties; the social democratic Labour Party (AP), the Socialist Left Party (SV) and the Centre Party (SP). AP is the biggest and most influential. The coalition is called the red-green coalition.

<sup>22</sup> The concerns about the use of the mapping tests have been raised in discussions with teachers and headmasters.

How well are young adults prepared to meet the challenges of the future? Are they all able to analyse, reason and communicate their ideas effectively? Do they have the capacity to continue learning through-out life? Parents, students, the public and those who run the education system need to know (OECD 1999: 7).

In other words, life-long learning and accountability are key issues. Moreover national unit of analysis is not enough: “Comparative international analyses can extend and enrich the national picture by establishing the levels of performance being achieved by students in other countries and by providing a larger context within which to interpret national results” (ibid.). It is further argued that there is a need for governments and the public to get “solid and internationally comparable evidence of educational outcomes” (ibid.) as a basis for policy decisions.

PISA is not directly linked to the curriculum as the tests of IEA, but to mastery. It “assesses how far students near the end of compulsory education have acquired some of the knowledge and skills that are essential for full participation in society”<sup>23</sup>.

Although the dominance of reading literacy, mathematical literacy and scientific literacy correspond to school subjects, the OECD assessments will not primarily examine how well students have mastered the specific curriculum content. Rather their aim is at assessing the extent to which young people have acquired the wider knowledge and skills in these domains that they will need in adult life. (OECD 1999: 9)

Moreover the mastery is connected to the development of human capital. “The emphasis on testing in terms of mastery of broad concepts is particularly significant in the light of the concern among nations to develop human capital” (ibid.:11).

The goal, as mentioned in 2.2.3, was to establish a set of indicators, and according OECD (1999: 10), PISA will provide the following:

- Basic indicators providing a baseline profile of the knowledge and skills of students;
- Contextual indicators, showing how such skills relate to important demographic, social, economic and educational variables;
- Indicators on trends that will emerge from on-going, cyclical nature of the data collection and that will show changes in outcome levels, changes in outcome distributions, changes in relationships between student-level and school-level background variables and outcomes over time.

---

<sup>23</sup> [http://www.oecd.org/pages/0,3417,en\\_32252351\\_32235918\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/pages/0,3417,en_32252351_32235918_1_1_1_1_1,00.html) (accessed 06.01.09).



### 2.3.3 Some technical aspects of PISA

PISA gathers data from cognitive assessments of reading, mathematics and science literacy every three years. The first assessment was conducted in 2000. Then the main focus was reading, in 2003 it was mathematics and in 2006 science. The next cycle will start in 2009, and the main focus will again be reading. In addition to the cognitive test items, in 2003 some items were designed to measure the problem solving abilities of students, and in 2006 items were added to test attitudes towards science. The students get two hours to answer the assessment, and additional 30 minutes to answer a questionnaire about social background, educational environment and learning habits. A questionnaire is also given to the school administration.

Each time an increasing number of countries have taken part in the study. In 2006 more than 400.000<sup>24</sup> 15 years old students from 57 countries participated. 67 countries have registered for 2009<sup>25</sup>. “Between 4500 and 10000 students will typically be tested in each country, providing a good sampling base from which to break down the results according to a range of students characteristics” (OECD 1999: 9).

The OECD secretariat is responsible for PISA, while the study is administrated by a *PISA Governing Board*, consisting of governmental representatives from participating countries<sup>26</sup>. There are two international expert groups; *Subject Matter Expert Group* and *PISA Questionnaire Expert Group*<sup>27</sup>, and a consortium lead by *Australian Council for Educational Research* (ACER) which is responsible for the design and implementation of the study. From the start, partners within the consortium were *National Institute for Educational*

---

<sup>24</sup> <http://www.oecd.org/dataoecd/15/13/39725224.pdf> (accessed 04.04.09). Se appendix D for more information.

<sup>25</sup> [http://www.pisa.oecd.org/pages/0,3417,en\\_32252351\\_32236225\\_1\\_1\\_1\\_1\\_1,00.html](http://www.pisa.oecd.org/pages/0,3417,en_32252351_32236225_1_1_1_1_1,00.html) (accessed 04.04.09).

<sup>26</sup> Anne-Berit Kavli, Head of Department, Norwegian Directorate for Education and Training is the Norwegian representative:  
[http://www.pisa.oecd.org/document/53/0,2340,en\\_32252351\\_32236359\\_33614005\\_1\\_1\\_1\\_1,00.html#Norway](http://www.pisa.oecd.org/document/53/0,2340,en_32252351_32236359_33614005_1_1_1_1,00.html#Norway) (accessed 04.04.09).

<sup>27</sup> The different organisation levels of PISA are found here:  
[http://www.oecd.org/document/53/0,3343,en\\_32252351\\_32235731\\_38262901\\_1\\_1\\_1\\_1,00.html#Who's Who in PISA](http://www.oecd.org/document/53/0,3343,en_32252351_32235731_38262901_1_1_1_1,00.html#Who's Who in PISA) (accessed 04.04.09). According to OECD (1999: 78) Svein Lie from University of Oslo was a science expert group member when the first PISA assessment was planned. According to OECD (2009: 416), he is still a member of the group. In addition Øystein Gutterud, Marit Kjærnsli and Rolf V. Olsen from UiO were consultants while preparing the science literacy test items to be used in PISA 2006.

*Measurement* (the Netherlands), Service de *pédagogie expérimentale de L'Université de Liège* and WESTAT<sup>28</sup> (OECD 1999:3). In 2006 the consortium was extended, and “[t]est development teams were established in five cultural-divided and well-known institutions namely ACER, CITO (The Netherlands), ILS (University of Oslo, Norway)<sup>29</sup>, IPN (University of Kiel, Germany) and NIER (Japan)” (PISA 2009: 29). The same year, ILS established *Unit for Quantitative Analyses of Education* (EKVA)<sup>30</sup>. Every country appoints a National Project Manager.

PISA is designed to be dynamic: “Given the long horizon of the project and the different relative emphases that will be given to the domains in each cycle, the OECD/PISA assessment frameworks clearly represent an instrument that will evolve” (OECD 1999: 16). Nonetheless, the main ideas and structures are the same from study to study. To illustrate how the cognitive test items are prepared and composed, I will use *PISA 2006 Technical Report* (OECD 2009).

Although each test has a main focus, all the literacies are tested every time PISA is conducted. When preparing for PISA 2006, the consortium therefore had to prepare items in all three literacies. The expert groups created 86 new test items in science, while 22 from 2003 were recycled, as were all the reading and mathematics items. The different items were divided in thirteen clusters<sup>31</sup>, seven containing science literacy test items, two containing reading and four containing mathematics literacy items. The clusters were combined in different ways in thirteen booklets. Each contained four clusters and each student got one booklet. The cognitive items were of two main formats; *multiple choice* and *free-response* items. The multiple choice items were of two types; the standard, where students got four possible answers to choose among, and the complex, where students had to respond to several statements with a yes/no, correct/incorrect etc. 59 test items were of the former, and 39 of the

---

<sup>28</sup> “WESTAT is an employee-owned corporation providing research services to agencies of the U.S. Government, as well as businesses, foundations, and state and local governments”: <http://www.westat.com/> (accessed 05.04.09).

<sup>29</sup> ILS is the abbreviation for *Institutt for Lærerutdanning og Skoleutvikling*, in English: *Department of Teacher Education and School Development*.

<sup>30</sup> In Norwegian: *Enhet for Kvantitative Utdanningsanalyser*, in abbreviation EKVA: <http://www.ekva.uio.no/english.html> (accessed 10.11.08).

<sup>31</sup> For an overview of the clustering of test items in PISA 2006, see OECD (2009: 29).

---

latter type. The free-response items were also of two types, the short closed - constructed response items and the open - constructed response item. The first type require a short sentence, a word or a number as an answer and constituted 18 test items, while the latter require a more extensive answer and often an explanation. This type constituted 71 test items.

## 2.4 Summing up

In this chapter I have given a short overview of the history of standardised test and explained my choice to use PISA as an example in the discussion of standardised tests. Further I have presented the arguments for PISA as they are explained in the first OECD/PISA document, and some technicalities connected to the assessment. The information given here will be used in chapter V and VI where the production and use of PISA as well as the reason for the development of a test culture will be analysed. However, first I will present the theoretical framework as well as the historical and economic context which I intend to use during my analysis in chapter V and VI.

### 3. The dialectics within and the contradictions between

The present exists because of the past. The development within education in general and of standardised testing in particular is, as I see it, the result of the “struggle” between contradictory but interrelated political, ideological, epistemological, ontological and methodological standpoints. Moreover, both the “struggle” and the outcome of the struggle are related to our social reality and social processes within it. Thus, to be able to answer my research questions I have to discuss the contradictions which have guided the choices that have been taken.

In this chapter I will therefore review the dialectics within and the contradictions between theories and thoughts driving the development in focus of my research questions. I will look into some grand theories, some middle range theories and some empirical findings that constitute the frame within which knowledge is constructed and educational choices are taken. First I will discuss ideological and theoretical positions, then I will identify and describe the driving forces of the current economic and political development and finally I will discuss the effects of the same development. My aim is to construct a theoretical framework that I can use while analysing my data in chapter V and VI. However, while working with my thesis, I have realised the connections that exist between my topic, my own methodological choices and the role of the comparative researcher in general, thus part of this chapter also will create a backdrop for chapter IV.

#### 3.1 Identification of the guiding theories

How do I look at the world? I have already said something by pointing at the contextual way in which I see the development within the field of education, and thus placed myself in a *dialectic* tradition. However there are more positions to make clear. Where do I place myself within the major theoretical influences in the field of comparative education; within *structural functionalism*, *Marxism*, *postmodernism* or *poststructuralism*? Am I sceptical to the direction of the development and want to contribute to the change of it, and hence confess myself to *critical theory*? Or do I think the challenges can be solved through the

*problem solving theory*? Moreover, what about my *epistemological*, *ontological* and *methodological* positions? The answers to those questions are crucial to all sorts of choices I have to take during the work with my thesis.

### 3.1.1 Marxism vs. Structural Functionalism

According to Kubow & Fossum (2007: 32) the major theories to use in comparative education are those classified as *modernist theories* and those found under the term *postmodernist* and *poststructuralist theories*. Due to the interconnections mentioned in the introduction of this chapter, those theories are also interesting for looking into the development of standardised testing, as well as when choosing my own research strategy.

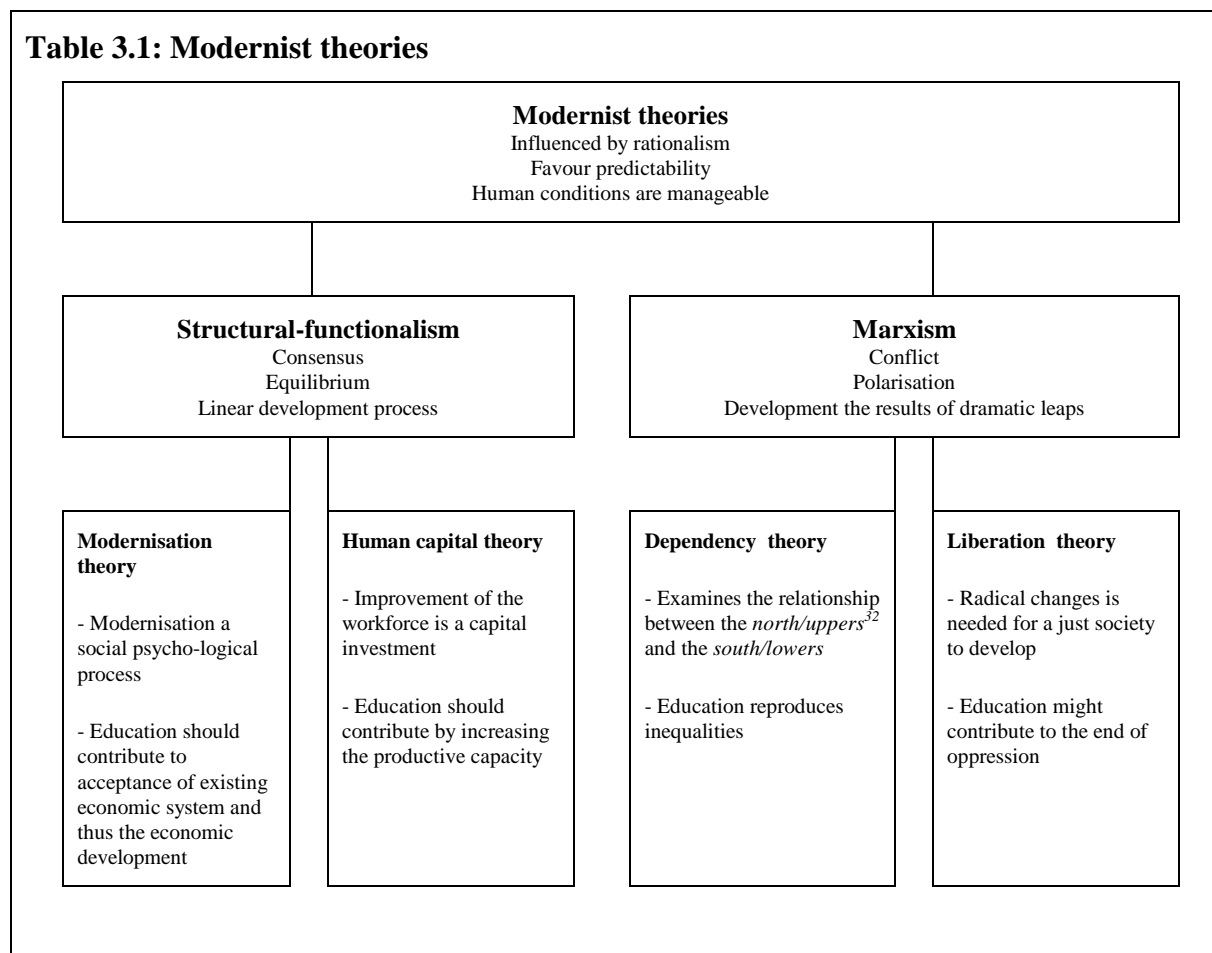
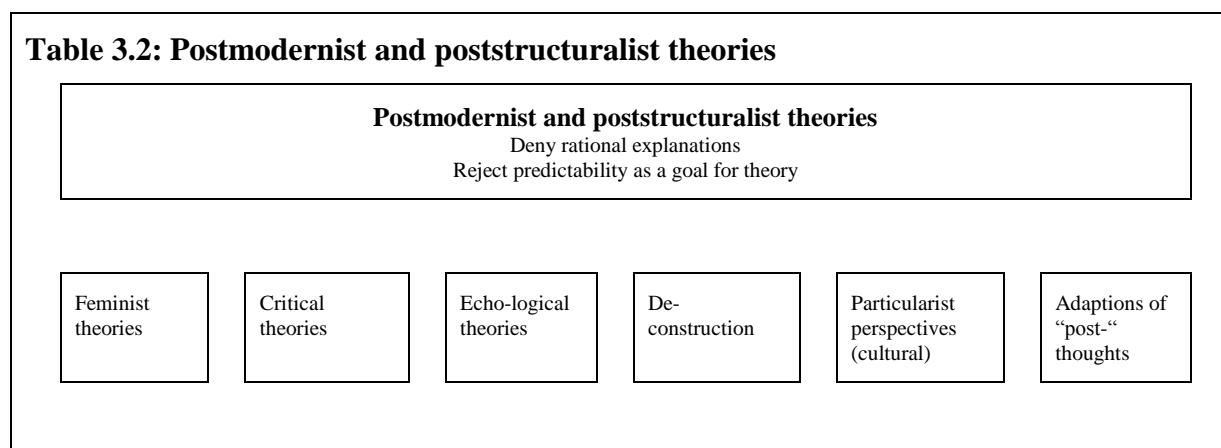
To help my analysis, I have made two simplified, schematic frames built on figures 2.1, 2.3 and 2.4 in Kubow & Fossum (2007: 32, 44, 45). Table 3.1 shows the modernist theories.

Kubow & Fossum (2007:33) state that “[e]ach of these major social theories makes contrasting claims, and each promotes different visions about the role of schooling within society, the ways instruction is designed and delivered, and the nature of relationships between school and society”.

There are obvious contradictions between the *modernist* and the *postmodernist* and *post structuralist* theories as shown in the tables on next page, however the term “post” has to be seen as something “trailing behind and coming in the wake of” (ibid.: 58); something following an earlier tradition. Hence the single category *critical theory* in table 3.2 hides four different schools of thoughts. Except for the last option, the relationship to the modernist theories is obvious. The four main schools of thoughts are the:

- neo-Marxist;
- poststructuralist;
- postmodernist;
- and the genealogical tradition (Kincheloe & McLaren 2005: 305).

Later in this chapter I will explain critical theory, why this theory is my choice and which of the schools of thoughts I have chosen.

**Table 3.1: Modernist theories****Table 3.2: Postmodernist and poststructuralist theories**

<sup>32</sup>The *north* and *south*, the *uppers* and the *lowers* are explained in 1.5.3.

### 3.1.2 Dialectics

“At the heart of dialectics is the idea that all ‘things’ are actually processes, that these processes are in constant motion, or development, and that this development is driven by the tension created by two interrelated opposites acting in contradiction with each other” Au (2007: 2 ). According to Au, who leans on a manifold of thinkers in his definition of dialectics (Allmann 1999, Gadotti, 1996, Ollman, 2003, Sayers, 1990, Woods & Grant, 2002), the opposites are deeply integrated. They require and depend on each other and constitute a unified whole. Dialectics also look at the world as a layered totality that is made of processes and chain of relationship.

To me educational development is the process occurring when contradictory but interrelated political, ideological, epistemological, ontological and methodological standpoints meet. This development does not happen by chance, but is the result of processes guided by conscious and contextual choices. To be able to discuss the contradictions that have guided those choices, I have chosen to use a dialectic approach in my thesis, because “[d]ialectics is the study of internal contradictions (...) [and the] development of contradictions is the driving forces of change (Kvale1996: 55)”. Kvale (1006: 57) deepen his argumentation by referring to Haug (1978):

If social reality is in itself contradictory, the task of social science is to investigate the real contradictions of the social situation and posit them against each other. In other words, if social processes are essentially contradictory, then empirical methods based on an exclusion of contradictions will be invalid for uncovering a contradictory social reality.

To understand the ongoing, dialectical processes; *status nascendi*, is more important than to describe the present; *status quo*. Kvale (1996:56) formulates it like this:

In dialectical thought there is an emphasis upon the new, what is under development. With a conception of the social world as being developed through contradictions, it is important to uncover the new developmental tendencies in order to obtain true knowledge of the social world. The statistical average or the representative case of the *status quo* is less important than the new tendencies developing as the *status nascendi*.

### 3.1.3 Critical theory

Critical theory is not one theory, but consists of four quite contradictory schools of thought (cf. 3.1.1.). They represent quite different theoretical positions. Therefore it is not easy to describe what critical theory is. Here I use Kincheloe & Steinberg (1997 in Kincheloe & McLaran 2005: 304) to define the concept:

We are defining a criticalist as a researcher or theorist who attempts to use her or his work as a form of social or cultural criticism and who accepts certain basic assumptions:

- all thought is fundamentally mediated by power relations that are social and historically constituted;
- facts can never be isolated from the domain of values or removed from some form of ideological inscriptions;
- the relationship between concept and object and between signifier and signified is never stable or fixed and is often mediated by the social relations of capitalist production and consumption;
- language is central to the formation of subjectivity (conscious and unconscious awareness);
- certain groups in any society and particular societies are privileged over others, although the reasons for this privileging may vary widely, the oppression that characterizes contemporary societies is most forcefully reproduced when subordinates accept their social status as natural, necessary, or inevitable;
- oppression has many faces and that focusing on only one at the expense of others (e.g. class oppression versus racism) often elides the interconnections among them;
- mainstream research practices are generally, although most often unwittingly, implicated in the reproduction of systems of class, race and gender oppression<sup>33</sup>.

These assumptions fit my own thoughts. To the second last I have to tie a comment, though. I agree about the many faces of oppression and their interconnections, however I regard one form more central than the others; class oppression. Nevertheless, after I found these assumptions, I knew I was a criticalist and that I wanted to use critical theory as a point of departure in my thesis.

---

<sup>33</sup> The organisation of the quotation in ballpoints is done by me.



Among the four schools of thought within critical theory, I position myself within the neo-Marxist tradition. This is due to my belief in *dialectical materialism* that “involves the fundamental assumption that the contradictions of material and economic life are the basis of social relations and of consciousness” (Kvale 1996: 55).

#### *Critical theory vs. problem solving theory*

In the introduction I mentioned the distinction between *critical theory* and the *problem solving theory*. Dale (2005) discusses the distinction between those two theories and what kind of consequences the choice of each of them might have in comparative studies. An extremely shortened and paraphrased version of Dale’s (2005:139) arguments are that the problem solving theory operates on the surface, values the individual over the structures of society, is disconnected to action and historical development and is meant to be used in decision making. On the other hand he argues that critical theory helps us to see the relations between the structures of society and the individual that actually shape the individual. Critical theory is connected to action and historical development and is meant to be used in concept formation. My interpretation of his analysis is that the problem solving theory is more *individualistic* and more *descriptive* and operates within status quo, and that the critical theory is more *holistic* and *explaining* and operates within an understanding of status nascendi.

### 3.1.4 The production of knowledge

What is “true” knowledge – and how it is constructed? These are *epistemological*, *ontological* and *methodological* questions interesting for my own choice of research strategy as well as for my topic.

Epistemology is seen as the philosophical theory of knowledge. Bryman (2004: 11) describes the concept like this: “An epistemological issue concerns the question of what is (or should be) regarded as acceptable knowledge in a discipline”. Is this “acceptable” knowledge derived from natural science methods or is it knowledge derived from a contradictory position? The first alternative reflects a *positivist* tradition, a tradition largely contextually independent and where the influence of the researcher should be as minimal as possible. The second alternative is built on a contextual tradition and is called *interpretivism*. To be short, a *positivistic* position is more *deductive*, “top down”, and an interpretivistic position is more *inductive*, “bottom up”. In between, we find *realism*. This position combines the positivistic

belief in the natural science methods, and the interpretivistic commitment to the external reality (ibid.).

Ontology tells us about the position of social phenomena and social actors. As in epistemology, there is more than one position. The *objectivists* assert that “social phenomena and their meanings have an existence that is independent of social actors” (Bryman 2004: 541). The *constructivists* assert that “social phenomena and their meanings are continually being accomplished by social actors” (Bryman 2004:538). The first position is more deductive and less contextual than the other. There is also a *realist* version of objectivism that “implies that an objective reality exists independently of the observer and that only one correct view can be taken of it” (Kvale1996: 66).

Methodologically it is the *qualitative* and *quantitative* approaches that hold the differential positions:

Qualitative research usually emphasizes words rather than quantification in the collection and analysis of data. As a research strategy it is inductivist, constructionist, and interpretivist. (...) Quantitative research usually emphasizes quantification in the collection of data. As a research strategy it is deductivist and objectivist and incorporates a natural science model of research process (in particular, one influenced by positivism) (Bryman 2004: 542).

Bryman adds that neither do qualitative nor quantitative researchers always subscribe to all features. Nowadays many researchers use methodological triangulation, which is a mix of qualitative and quantitative methods. Among others is Patton (2002), who supports methodological freedom and the use of mixed methods and methodological triangulation. Nevertheless, in chapter VI I will ask whether the present educational development gives precedence to one methodology over the other. In chapter IV I will define my own methodological choice.

### 3.1.5 What kind of qualifications are “needed”?

What is schooling about? What should schooling “produce”? Should it give *Bildung* to those enrolled or just produce suitable workforce? Is human capital meant to be a capital for the persons that hold it, or is it meant to be a capital suitable to sell? How do the raise in the amount of standardise testing influence what the school “produce”? Masuch (1973 in Brock-Utne 2007a), has constructed a *qualification framework* that is suitable as a tool when some

of those questions are discussed in chapter V. I have designed a simplified, graphical version of the framework on the basis of the English translation of Brock-Utne (2007a)<sup>34</sup>.

<b>Table 3.3: Qualification framework</b>			
<b>Main groups</b>	<b>Qualifications necessary to / attached to</b>	<b>Subgroups</b>	<b>Characterized by</b>
<i>Proficiency qualifications</i>	- carry out a given work process.	<i>General proficiency qualifications</i>	- general skills like reading. Independent of vocational function. The more industrialized a country is the more need for such qualifications.
		<i>Special proficiency qualifications</i>	- special skills needed for a vocation. The more industrialized a country is the shorter the lifetime of such skills.
<i>Adaptability qualifications</i>	- attitudes and ways of behaving.	<i>Active adaptability qualifications</i>	- diligence and perseverance. Help to carry out a work process with great intensity.
		<i>Directly adaptability qualifications</i>	- obedience and a sense of duty. Help to carry out a work process with willingness and the best ability.
		<i>Indirectly adaptability qualifications</i>	- indifference and apathy. Prevent rebellion behaviour.
<i>Creative qualifications</i>	- the development of productive forces needed for capital accumulation.		- critical sense, independence, openness, the ability to synthesize and develop new insight, the ability to enter into constructive cooperation.

## 3.2 Identification of driving forces

What is the context within which present economic development happens? What have been the driving forces behind and what have been the impacts of the development? Those questions are central to my analysis in chapter V, thus I need to define some central concepts, identify the driving forces connected to and discuss the effects of the economic development at global and public level.

<sup>34</sup> The original is written in German, and my German does not hold an academic level.

### 3.2.1 Globalisation

Globalisation as a concept got its breakthrough in the early 1980s and was used to explain economic and technological changes. In Norway the concept was used by politicians and industry leaders to explain why the industry had to get other frames to be able to compete internationally (Eriksen 2006: 22). From the early 1990s the use of the concept has boomed. Inspired by Eriksen I googled the word written with “z” and got 106 million hits, when using an “s” I got nearly 40 million. The concept globalisation “has come from nowhere to be almost everywhere” Giddens (1999) argues (in Crossley and Watson 2006: 53). However, what does the concept mean? Crossley and Watson (2006: 53) emphasise the difficulties in defining it by stating that globalisation is “a complex and highly contested term – and one that is widely used but open to multiple interpretations”, and “...there is a wide disagreement relating to the origins, mechanisms, significance and implications of the concept”. Eriksen (2006: 21) argues that it is a concept that might mean a lot, very little or different things at the same time. At page 57 he deepens his argument by explaining the different aspects of globalisation as he sees it: From an economic aspect, we can talk about *global capitalism* and from an ideological aspect, *neoliberalism*. Other aspects or results of globalisation are tax *havens*, *launder*, *international drug dealing* and *fraud*<sup>35</sup>, which he suggests to give the joint designation *crime*, still others are *polarising*, *marginalisation* and *imperialism*. In conclusion he simply ends up renaming the concept globalisation by calling it *global apartheid*.

This is a rather challenging way to look at a concept that for more or less 25 years has been used as something that simply *is*, something impossible to escape, a natural force! However, I am fully aware of the fact that people look at the concept differently. In their book, Crossley and Watson (2006: 53 – 55) use a classification made by Held *et al.* (1999) of three main tendencies of thought. *The hyperglobalists* promote globalisation and view it as *the* way to develop the world. Moreover they predict the weakening of the nation-state and breakdown of the traditional *north-south* divide. *The sceptics* criticise globalisation as a neo-imperialistic and devastating development track and for the power it gives to the corporate culture of the *World Trade Organisation* (WTO) and other organisations. They also argue against the deterministic view of those who promote and support globalisation. The

---

<sup>35</sup> He uses the example of what happened in Enron while explaining what he means by fraud. Post 2008 there is no problem to find other examples!

---

alterglobalisation movement<sup>36</sup> is found in this category. *The transformationalists* are not promoting globalisation, but accept it and adapt to it. They do not have a future scenario to the kind of world globalisation develops. One well known scholar and one of Tony Blair's advisors, Anthony Giddens, argues in a direction that places him in this third category (Giddens 1998). He and others sharing the same opinion as well as representatives from the hyperglobalists would never make an analysis like the one of Eriksen (2006). However, I use the analysis of Eriksen because I agree with his position. Moreover it makes the complexity of the concept clear. For me the complexity is of such a kind, that the concept can not do anything but blur an analysis. Hence, I will not use the concept globalisation in my paper, unless the formulation is used by someone in a quotation. I will simply call a spade a spade, and use the term global capitalism when writing about the economic policy, neoliberalism when referring to the driving ideology, and marketisation of education when I talk about what happens to education in a globalised world. Moreover I will use the term neoliberalisation, a term taken from Harvey (2005), while describing the process of so called globalisation.

### 3.2.2 Neoliberalism

As mentioned in chapter I, the marketisation of education is linked to the concept neoliberalism. Here I will explain the concept in a theoretical and historical perspective. Later in this chapter I will connect it to the global development as well as to the development within the public sector.

#### *The rise and fall of economic liberalism*

Adam Smith formulated the theory of economic liberalism in the 18<sup>th</sup> century. His ideas gained momentum in the 19<sup>th</sup> century during a period of growth and rising world trade (Marsdal & Wold 2004). However the growth stagnated, and the world experienced the big economic crisis in 1929 that was followed by WWII. When 44 countries gathered in Bretton

---

<sup>36</sup> The Social Movement consists of organisations (*Non Governmental Organisations* (NGOs), Unions, other) and private persons from all over the world. They have been organising the struggle against the current form of globalisation since the 1990s. The movement has been organising alternative meetings during the WTO negotiations and their first big victory was Seattle in 1999, when they managed to stop the negotiations. They also arrange World Social Forums parallel to World Economic Forum, and in January 2009 the ninth was arranged in Belem, Brazil. Also regional, national, and locale forums are organised regularly. Their slogan is: *Another world is possible!* Believing in another type of globalisation, they call themselves alterglobalists.

Woods in 1944, it was to create an economic world order to prevent a repetition of 1929 (Harvey 2005). Though the conference did not follow the more radical proposals from John Maynard Keynes (Garbo 2008), the rules that were agreed upon made it possible for the active state to put up restrictions to protect the society from the free market forces and to curb capitalism (Harvey 2005, Marsdal & Wold 2004). After WWII the world experienced the greatest economic growth ever and the redistribution of wealth and power from the elites toward the common was considerable (Harvey 2005). Harvey calls this period *embedded liberalism* “to signal how market processes and entrepreneurial and corporate activities were surrounded by a web of social and political constraints and a regulatory environment” (ibid:11).

### *The creation of neoliberalism*

However, around 1970 the economic decline started; the inflation, the oil prices and the unemployment increased, the profit declined and the core capitalist countries entered into an accumulation crisis (Eriksen 2006). There was a strong polarized debate about how to solve the crisis. On one side we found those wishing a strong state and central planning; the *left* and the social democrats. On the other side was the *right* represented by business power wishing market freedom. In 1990s the *right*, also called the Washington Consensus, won (Harvey 2005). How did it happen?

Harvey (2005) emphasis two things while explaining why the *right* side won; the weakening of the *left*, particularly after the fall of the Berlin Wall<sup>37</sup> and the following breaking down of USSR, and the establishing of neo liberal think tanks.

Those sceptical to the class compromises during the embedded liberalism, such as Friedeich von Hayek, Milton Friedman<sup>38</sup> and their supporters, had since 1947 consciously worked to counteract the development. To succeed in this work, Friedman argued for the need to erase the ideas of the political *left* and the unions through the upbringing and training of people with another world view; the liberalistic world view (Marsdal & Wold 2004). From the early 1970s, think tanks to re-establish economic liberation popped up particularly in the US and

---

<sup>37</sup> The fall of the Berlin Wall in 1989 is used as a symbol of the breaking down of the USSR as a super power.

<sup>38</sup> In 1974 they got the Nobel Prize in economic. According to Harvey (2005: 22), this offered academic respectability to the neoliberal theory.

---

the UK. They were heavily supported by the business sector, and their goal was to construct consent for a liberal policy among people; to influence the way people think. “For any way of thoughts to become dominant, a conceptual apparatus has to be advanced that appeals to our institutions and instincts, to our values and our desires, as well as to the possibilities inherent in the social world we inhabit” (Harvey 2005: 5).

The new conceptual framework was introduced at the universities and in schools, in the publishing houses, the media and in the court. By playing on the concepts of Freedom, Choice and Rights, on traditional values and sometimes also on phobia against the state as a suppressing and even a communist tool, they managed to get support. During the elections in 1979 and 1980, respectively Margaret Thatcher and Ronald Reagan came to power in the UK and the US. After 10 years in power, Thatcher greeted Hayek on his 90 years birthday, expressing gratitude for his decisive work (Crockett 1994 in Marsdal and Wold 2004: 49). They had a common understanding of the importance of changing the way people think in order to change the world. “Economics are the method”, she [Thatcher] said, “but the object is to change the soul” (Yergin & Stanislaw 1999 in Harvey 2005:23).

Over time the economic understanding taught by many of the most prominent universities of the US, such as Harvard and Stanford, travelled the world with the scholars they produced. The neoliberal thinking crept into the administrations of different countries and international institutions like *International Monetary Fund* (IMF), the *World Bank* and the WTO. “[B]y 1990 or so most economic department in the major research universities as well as the business schools were dominated by neoliberal modes of thought. The importance of this should not be underestimated<sup>39</sup>” (Harvey 2005: 54). On page 3 Harvey (2005) deepens his position: “Neoliberalism has, in short, become hegemonic as a mode of discourse”.

Both the weakening of the *left* and the work of the think tanks gradually changed the social democratic way of thinking and turned them into neoliberal followers in their own countries.

---

<sup>39</sup> His analysis of the role of the financial sector in the US was supported by the report *Sold Out: How Wall Street and Washington Betrayed America* that came out 04.03.09. The report is written by *Wall Street Watch* and reveals that the financial sector has used huge amount of money to pay lobbyists, finance think tanks and election campaigns for politicians since the time of Reagan. The last 10 years they have used more than \$5 billion. They efforts manage to deregulate the economy, something that paved the way for the financial crisis in 2008, according to the report. The whole report is found at: [http://wallstreetwatch.org/reports/sold\\_out.pdf](http://wallstreetwatch.org/reports/sold_out.pdf) (accessed 14.03.09).

Over the time more and more steps, small and big, were taken in a neoliberal direction all over the world. An early step came in 1971, when the gold standard was removed and by this the fixed exchange rate. Over time, most of the restrictions disappeared. The redistribution of wealth and power reversed, and the restoration of class power started (ibid.).

### *The goals and keywords of neoliberalism*

Neoliberalism is in the first instance a theory of political economic practices that proposes that human well-being can best be advanced by liberating individual entrepreneurial freedoms and skills within an institutional framework characterized by strong private property rights, free markets, and free trade. The role of the state is to create and preserve an institutional framework appropriate for such practices. (...) [I]f markets do not exist (in areas such as land, water, education, health care, social security, or environmental pollution) then they must be created (Harvey 2005: 2).

As seen in the quotation, one goal is well-being. Others are to eliminate bureaucracy and reduce cost, to improve efficiency, productivity and quality. Keywords are individual freedom of action, expression and choice, especially the freedom of business in a free market, private enterprise, entrepreneurial initiative and intellectual property rights. What is positive is competition, privatisation, deregulation and flexibility, what hinders the development is the lack of private property rights and interventions into the market. Such policy is supposed to carry a trickle down effect, and is thus seen as the best way to eliminate poverty. What the state is supposed to do is to create a framework for the policy to work, but not intervene (Harvey 2005). Later in this chapter I will look into the impact of this policy and whether there is coherence between the rhetoric and the reality. But first another concept must be mentioned.

### *The nurses of neoliberalisation*

The neoliberalisation process of the world differs in time, place and strength. In Chile the coup in 1973, in UK the election of Thatcher in 1979, in China decisions during the leadership of Deng Xiaoping from the late 1970s<sup>40</sup>, in US the election of Reagan in 1980, in the south the *Economic Structural Adjustment Programme* (ESAP)<sup>41</sup> of the 1980s and in

---

<sup>40</sup> Although the development in a neoliberal direction in China has been much more controlled by the state (Harvey 2005).

<sup>41</sup> The name of the policy of the IMF/the World Bank has changed over the years. It changed from ESAP to *Structural Adjustment Programme* (SAP) and then to *Poverty Reduction Strategy Papers* (PRSP). According to Steven Klees in his key note speech at *The Implications of Language for Peace and Development Conference* (IMPLAN) held in Oslo 2-3 May 2008, the policy remained the same in spite of the different names.



---

former Eastern Europe the break down of the USSR in the early 1990s were important turning points (Harvey 2005).

In some countries such as the US, the UK, Chile, Australia and New Zealand the neoliberalisation process started early and were extensive. In the majority of the countries, however, the ideas were only partly implemented. In most of the countries democratically elected governments of all political colours have been leading the process , but some countries, mainly in the *south*, were more or less forced to open up to the neoliberal ideas. I have used the analysis of Harvey (2005) to describe the *midwives* of neoliberalisation, but who where the *nurses*?

Since WWII there has been a tremendous growth in the number of national, bilateral, intergovernmental, nongovernmental, quasigovernmental, religious and philanthropic organisations (Amin 2006, Arnove 2003, King 2007, Samoff 2003). Over the time they joined with professionals and scientists in systems of like-minded actors of global engagement and transnational influence. Jones (2007) speaks about the emergence of *epistemic communities*. According to him, especially the communities grounded in the economy, business and financial sectors have been powerful. Their like-minded way of thinking has been promoting the process of neoliberalism, and their influence has been tremendous both in *south* and *north*. Parallel to this development, the states slowly lost some of their independence and became more interdependent. Some of the most influential organisations have been the IMF, the World Bank, the WTO, the OECD and *European Union* (EU). In chapter V I will discuss their impact on education.

### 3.2.3 Neoconservatism

As already mentioned, liberalism gave the moral legitimacy to capitalism. For a period of 25 years or so after WWII, embedded liberalism guided the economic policy of the western world, also called the golden age of capitalism (Eriksen 2006). Later, and especially after the fall of the breakdown of USSR, neoliberalism has guided a capitalistic policy that has been a

---

leading policy in more and more countries (Eriksen 2006, Harvey 2005). When neoliberalists are neoliberals, the capitalist class of today can be both neoliberal and neoconservative.

Neoconservatists agree with the ideas of the neoliberal, however

it veers away from the principles of pure neoliberalism and has reshaped neoliberal practices in two fundamental respects: first, in its concern for order as an answer to the chaos of individual interests, and second, in its concern for the overweening morality as the necessary social glue to keep the body politic secure in the face of external and internal dangers (Harvey 2005:82).

This need for order and morality is highly visible in the present education policy. Apple (2006) is highly concerned about this point. He also argues that we must not forget the role of the *Authoritarian Populism* and the *Professional and Managerial New Middle Class* while discussing the current development. I will return to this point in chapter V when discussing the counteracting development trends of decentralisation and centralisation.

### 3.3 The effects of the neoliberalisation process

There are lots of studies and books written on how the combined policy of the neoliberals and the neoconservatives function. Within the scope of this text I am unable to pay respect to all the different view-points. In this section I have merely picked some official statistics and some critical analysis in order to view the paradoxes between the rhetoric and the reality of this policy. My questions are: Did the policy deliver what it promised? Did it raise the growth, erase the poverty and give freedom to the people?

#### 3.3.1 The global effects

Aggregate global growth has declined remarkably (Chomsky 2006, Harvey 2005, Marsdal & Wold 2004). In the 1960s it was 3,5%, in the 1970s 2,4%, in 1980s 1,4%, in 1990s 1,1% and since 2000 it has barely reached 1% (Harvey 2005). However some have skimmed the cream. “The world’s 200 richest people more than doubled their net worth in the four years to 1998, to more than \$1 trillion. The assets of the top three billionaires more than combined GNP of all least developed countries and their 600 million people” (UNDP 1999: 3). The income gap between the richest fifth in the *north* and the poorest fifth in the *south* has risen.

In 1960 the gap was 30:1, in 1980 it was 41:1 and in 1989 it was 59:1, a tendency that is visualised by the *Champagne Glass* printed for the first time in the Human Developing Report of 1992 (UNDP 1992: 35). According to the report of 1999 (in Harvey 2005:19) the gap in 1997 was 74:1. 2008 numbers from the World Bank shows that the 10% of the richest people in the world, consume 59% of all there is consumed, whereas the 10% poorest only consumed 0,5<sup>42</sup>. If we look at the *Gini coefficient*<sup>43</sup>, the trend in the world as a whole shows a raising inequality between 1951 and 1999; from 0,345 in 1951 to 0,40 in 1999. The lowest inequality was according to the graph at the end of the 1960s. In Norway the Gini coefficient has raised from 0,24 in 1994, to 0,33 in 2005 (Wahl 2009: 132).

The inequality gap has widened both between and within countries, there is a *north* in the *south*, and a *south* in the *north*. Harvey (2005) mentions the growing inequality in China and in former Eastern Europe and also the same tendency in OECD countries since the 1980s. Norway is one of them (Wahl 2009). The distribution between the rich and the poor in the US is by now on the same level as it was before the crack in 1929, in other words we can see the tendency of redistribution of class power, according to Harvey (2005). “Redistributive effects and increasing social inequality have in fact been such a persistent feature of neoliberalization as to be regarded as structural to the whole project” (ibid.:19). Navarro (2004) argues that the growing inequalities are an evil in itself and that it has negative impact on health, quality of life and even age of death. The bigger the differences are within a country, the more difficult it is to be poor. Since the 1980s informal employment has risen tremendously and “all global indicators on health levels, life expectancy, infant mortality, and the like show losses rather than gains in well-being since the 1960s” (Harvey 2005: 154). His conclusions are quite harsh: “The reduction and control of inflation is the only systematic success neoliberalization can claim” (ibid.: 155) and “[t]he main substantive achievement of neoliberalization, however, has been to redistribute, rather than to generate, wealth and

---

<sup>42</sup> A bar graph dividing the worlds population in deciles and showing the distribution of consumption [http://siteresources.worldbank.org/DATASTATISTICS/Resources/WDI08\\_section1\\_intro.pdf](http://siteresources.worldbank.org/DATASTATISTICS/Resources/WDI08_section1_intro.pdf) : 4.

<sup>43</sup>The Gini coefficient: A number between 0 and 1 telling us about inequality within countries, the lower the number, the higher the equality. Inequalities are also measured in per cent, called Gini index. This index goes from 0 to 100, the lower the number, the lowest inequality. A graph showing the growing inequality in the world is found here: [http://www.unescap.org/tid/artnet/mtg/cb3\\_d2s2low.pdf](http://www.unescap.org/tid/artnet/mtg/cb3_d2s2low.pdf) : 7 (accessed 17.01.09).

income” (ibid.:159). One of his main issues is that the policy leads to accumulation of dispossession, through a reversed Robin Hood policy (See also Harvey 2003).

It seems that neoliberalism has certainly not been a success for the *lowers*. Critical thinkers, but also former central neoliberal actors like Georg Soros<sup>44</sup>, Joseph Stiglitz<sup>45</sup>, and other, have seen it for a while. “‘What a peculiar world’ Stiglitz quizzically remarks, ‘in which the poor countries are in effects subsidizing the riches’”(Stiglitz 2003 in Harvey 2005: 74). Another giving a verdict of neoliberalism is Alan Greenspan<sup>46</sup>. After the economic break down in the autumn of 2008 he stated “Market competition and free Markets did break down”<sup>47</sup>. What we have to bear in mind, however, is that he first recognised this when the system did not even work for the *uppers*. Yash Tandon’s<sup>48</sup> dry remark at Norway Social Forum November 7<sup>th</sup> 2008 is worth mentioning: “Your newspapers are full of descriptions about the economic crisis, but we [he represents the *south*] have experienced an economic crisis for 30 years!” Another thing to bear in mind is a real paradox. When the crisis hit in 2008, the states had no problems to intervene with the markets. They did so by pumping money into structures that might be seen as a part of the reason for the crisis; the bank system. This was not any surprise for a man like Harvey (2005: 48) who four years earlier argued “In the event of a conflict between integrity of financial institutions and bondholders’ returns, on the one hand, and the well-being of the citizens on the other, the former was to be privileged”. The other side of this story is the old slogan “What is good for Wall Street is good for Main Street”, but honestly; are there any reasons to believe in this trickle down effect any more? Does not the gap between promises and what we get, between the rhetoric and reality prove the weaknesses of the system?

---

<sup>44</sup> Soros is a global financier and a philanthropist; [http://www.soros.org/about/bios/a\\_soros](http://www.soros.org/about/bios/a_soros), (accessed 23.01.09).

<sup>45</sup> Stiglitz was a Chief Economist at the World Bank from 1996 until 1999. He became quite critical of the Bank’s policy and resigned <http://www.whirledbank.org/ourwords/stiglitz.html> (accessed 15.05.09).

<sup>46</sup> Alan Greenspan is the former Chairman of the Federal Reserve of the US from 1987 <http://www.sn1.no/.search?query=Greenspan%2C+Alan> (accessed 19.05.09) to 2006: [http://en.wikipedia.org/wiki/Alan\\_Greenspan](http://en.wikipedia.org/wiki/Alan_Greenspan) (accessed 23.01.09).

<sup>47</sup> This was said during a hearing in the congress and was broadcasted through the Norwegian Radio (NRK) 24.10.08.

<sup>48</sup> Yash Tandon was the Executive director of the South Centre and a doctor from London School of Economic. [http://www.southcentre.org/index.php?option=com\\_content&task=view&id=581&Itemid=144](http://www.southcentre.org/index.php?option=com_content&task=view&id=581&Itemid=144) (accessed 11.02.09).

### 3.3.2 The effects on public services

What drives the capitalistic system is the wish to make profit. When the profit is accumulated and the accumulated money does not find areas for new investments, a problem will emerge. From the 1970s this problem was “solved” mainly in two ways. One was to give huge loans to the *south* that within some years placed them in a bottomless depth crisis and further into the claws of IMF, the World Bank and ESAP (Bond 2006, Harvey 2005, SAPRIN 2005). Another was to try to enter into the public sector in the *north*.

In countries with strong trade unions that had fought for the workers right and the development of a strong security net, like the Scandinavian welfare states, such a development was highly controversial. Our system was solidly grounded in a social democratic way of thinking. To deregulate the economy, limit public spending and liberate the public sector by implementing market solutions that were needed to meet the neoliberal demands, was understood as a move “backwards” in the minds of the many. However, also here the belief in free market economy and free competition as the most effective instrument to create the foundation of solid economy and freedom in society had its supporters and over the time it was also partly implemented here (Harvey 2005, Jensen 2007, Lund 2008a & b).

#### *New Public Management (NPM)*

One concrete tool developed to promote the neoliberalisation of the public sector is NPM (Lund 2008a). This and related systems can be viewed as tools to facilitate structural adjustment both in the *south* and *north*. “NPM is the transfer of business and market principles and management techniques from the private into the public sector, symbiotic with and based on neo-liberal understanding of state and economy” (Drechsler 2005: 1).

According to Drechsler (2005) NPM as an ideology is on the defensive within the advanced public administration scholarship, especially in Europe. Chang (2002: 128) offers an explanation while arguing that “[t]he plain fact is that the Neo-Liberal ‘policy reforms’ have not been able to deliver their central promise – namely, economic growth”. Nonetheless, NPM and NPM- like models are still alive both in scholarship and policy all over of the

world, for example both in Denmark (Lund 2008 b) and Norway, where *Government Agency for Financial Management (SSØ)*<sup>49</sup> can be seen as an example (Jensen 2007, Wahl 2009).

During the last years critiques about the models have risen, as in the case of SSØ (Lie 2009: 8). One central question is whether the private and the public sectors are similar to such an extent that using the same systems, measures and vocabulary makes sense. To those in favour of neo-liberalism and NPM the answer most probably is “yes”. However, both well-informed lay people and critical economists emphasise on the differences between the spheres, of their different aim and legacy. While business focuses on making profit, the role of the state ideally should be to focus on the development of a system that includes and benefits all citizens. Drechsler (2005: 2) puts it this way:

The use of business techniques within the public sphere thus confuses the most basic requirements of any state, particularly of a Democracy, with a liability; regularity, transparency, and due process are simply much more important than low costs and speed. (...) It could be argued that most activities carried out by the public sector are there precisely because no direct profit or gain can be made.

Another central question is the costs that NPM and NPM like models impose on public sector. According to Reikvam *et. al.* (2008) it is not easy to calculate the exact costs of such models in Norway, but by using different reports, where Asplan Analyse (2005) is one, they conclude that the system is both expensive and bureaucratic. As a rough calculation they fix the extra costs of NPM-reforms in Norway at about 12 000 million Norwegian kroner (NOK), or 15 000 man-labour years. This is payment to private organisations, consultancies, advertisers etc. that is used to assist the public sector to implement the models imported from the private sector. This can not be anything else than what is called “a straw into the public purse”? Apple (2006) and Kohn (2000, 2002) describes the same tendency in the US, Hatcher (2002, 2009) and Jones *et.al.*(2008) in the UK.

For the education sector it is interesting to note that Norwegian Universities and University Colleges use half a billion NOK for advertisements to compete for students (Reikvam *et. al.* 2008). These costs are mainly expenses for the public and gain for some private public

---

<sup>49</sup> In Norwegian: Senter for Statlig Økonomistyring, (SSØ). This is a pilot project established by the *centre/right* government in 2004 and continued by the red-green coalition. Within the Norwegian Confederation of Trade Unions (LO), the opposition against SSØ is raising, something that will be one of the discussions at the LO congress May 2009 [http://www.sfso.no/templates/Page\\_135.aspx](http://www.sfso.no/templates/Page_135.aspx) (assessed 21.04.09).

---

relations companies. Thus, the introduction of competition in Higher Education has created a big market for public relations companies.

### 3.4 Summing up

In this chapter I have presented the theoretical framework and the economic and political context that will guide my analysis in chapters to come. In 3.1 grand theories and middle range theories were presented. In 3.2 I identified the driving forces behind the economic policy since the WWII, and in 3.3 I looked into the effects of this policy worldwide and within the public sector in the *north*. The effect of this policy on education will be discussed in the two main chapter of this study, chapter V and VI. Before turning to educational issues, however, I will present the methodological chapter.

## 4. A bricoleur

Research design is the “glue holding the research together”<sup>50</sup>. This chapter contains my choice of design, and will cover the choice of research strategy, research type, research method, data collection methods and analysis. It will also discuss the challenges I have met during the research process.

### 4.1 A qualitative research strategy

The departure point of this research is standardised testing. More concrete the league tables of PISA. Results from comparative quantitative research will thus be the starting point of my research. Nonetheless I have primarily chosen a qualitative research strategy for my research on the league tables, and I have eight reasons for this choice of approach.

*First* of all, I have always made patch works and collages, been a quilt maker in a wide understanding of the word. I have always torn yarn, fabrics, photos, graphic, ideas and thoughts in pieces and tried to put them together into a new whole. My idea was to do the same in this study, by using fragments from all my different undergraduate subjects, experiences and collected data, and make it into a whole. Hopefully a new whole, but this has proved to be difficult. To make nice, new quilts are always difficult. When I found the description of the qualitative researcher as a quilt maker or bricoleur in Denzin & Lincoln (2005) and Kincheloe & McLaren (2005), I decided to present myself as a *bricoleur*:

The qualitative researcher may be described using multiple and gendered images: scientists, naturalist, field-worker, journalist, social critic, artist, performer, jazz musician, filmmaker, quilt maker, essayist. The many methodological practices of qualitative research may be viewed as soft science, journalism, ethnography, bricolage, quilt making or montage. The researcher, in turn, may be seen as a bricoleur, as a maker of quilts, or, as in filmmaking, a person who assembles into montages (Denzin & Lincoln 2005: 4).

They present five different bricoleurs; the *interpretive*, *narrative*, *theoretical*, *political* and the *methodological*. At the present level of my understanding of being a qualitative researcher, I see myself as a combination of the theoretical, critical and political bricoleur:

---

<sup>50</sup>Halla Holmarsdottir during the advanced course in qualitative research methods – part 1, the fall of 2008.



- 
- The theoretical *bricoleur* reads widely and is knowledgeable about the many interpretive paradigms (feminism, Marxism, cultural studies, constructivism, queer theory) that can be brought to any particular problem (Denzin & Lincoln 2005: 6).
  - The political bricoleur knows that science is power, for all research findings have political implications. There is no value-free science. This researcher seeks a civic social science based on politics of hope (Denzin & Lincoln 2005: 6).
  - The critical bricoleur stresses the dialectical and hermeneutic nature of interdisciplinary inquiry, knowing that the boundaries that previously separated traditional disciplines no longer hold (Kincheloe 2001 in Denzin & Lincoln 2005: 6).

*Secondly* a qualitative approach opens up for a dynamic research suitable for a bricoleur; the research design is not as rigid as a quantitative one and allows the researcher to take part in a research process guided by her growing understanding of the field in question. Patton (2002: 40) calls this *emergent design flexibility* and describes it as “[o]penness to adapting inquiry as understanding deepens and/or situation changes”.

*Thirdly*, I do not have any theory to prove, but a problem or phenomenon to look into; hence my inquiry will be *inductive*.

*Fourthly*, I will conduct a *naturalistic inquiry* (Patton 2002: 40) in a real-world situation neither manipulated nor controlled by me.

*Fifthly*, a qualitative research allows *purposeful sampling*; “Cases for study (...) are selected because they are ‘information rich’ and illuminative, that is, they offer useful manifestation of the phenomenon of interest” (Patton 2002: 40).

*Sixthly*, to be able to find out about the role of standardised testing, I have to investigate the *process* in which the tests have played a role. As I see it, processes are difficult to investigate quantitatively.

The *seventh* reason for the choice of methodology has to do with my ontological standpoint. According to Prior (2008), not only humans and their doing enter into fields of action, also documents do<sup>51</sup>. Documents are produced and used by humans, and their content are formulated, interpreted and understood by them. That makes documents un-static agents in their own right situated in a field of action:

[E]ach and every document stands in a dual relation to the field of action. First, it enters the field as a receptacle (of instructions, commands, wishes, reports etc.). Secondly, it enters the field as

---

<sup>51</sup> Documents are my main data source, as will be showed later in this chapter.

an agent in its own right. And as an agent a document is open to manipulation by others: as an ally, as a resource for further action, as an enemy to be destroyed, or suppressed (Prior 2008: 3). She also states “[T]here is forever a dynamic to ‘the field’ in such a way that things, such as documents and the information they contain, can influence and structure human agents every bit as effectively as the agents influence the things” (Prior 2008:3). In other words, a document is never made or used, nor does it exist, in a vacuum. Moreover, a document has three different aspects to take into consideration; the production, the use and the content. This is a constructionist way of looking at documents. In the ontological tradition called objectivism, documents are regarded as more static and fixed, and emphasise is on the content of the document, not on the production and use of it.

My *last argument* as to why my main emphasise will be on qualitative approach is epistemologically connected and adheres to the discussion of the nature of qualitative and quantitative research. Are they “grounded in incompatible epistemological principles” or not (Bryman 2004: 54)? I will not enter into a deep discussion, thus I am not against quantitative research per se. However the PISA research that I have looked into fits very well to a positivistic approach to knowledge production (cf. chapter VI), an approach I question through the use of critical theory (cf. 3.1.3.). As a bricoleur I “believe in active human agency [and] refuses standardized modes of knowledge production” (Kincheloe & McLaren 2005: 317). Accordingly I am sceptical of the role that quantitative based knowledge has got within the field of education. Where the scepticism towards positivism and the use of the knowledge it produces meet, my own choice of research strategy and the topic of my thesis also meet. To distinguish between quantitative and qualitative research thus becomes important for my own analysis in chapter V and VI. Here I use the words of Denzin & Lincoln (2005: 10) to draw up the borders:

The word qualitative implies an emphasis on the qualities of entities and on processes and meanings that are not experimentally examined or measured (if measured at all) in terms of quantity, amount, intensity, or frequency. Qualitative researchers stress the socially constructed nature of reality, the intimate relationship between the researcher and what is studied, and the situational constraints that shape inquiry. Such researchers emphasize the value-laden nature of inquiry. They seek answers to questions that stress how social experience is created and given meaning. In contrast, quantitative studies emphasize the measurement and analysis of causal relationships between variables, not processes. Proponents of such studies claim that their work is done from within a value-free framework.

### 4.1.1 Research type

#### *Lack of consensus*

I got the concepts *research type*<sup>52</sup> and *research approach*<sup>53</sup> from Holmarsdottir who presented four main types of qualitative research types/approaches. Bryman (2004: 56) uses the concept *research design* when he talks about the same concepts, and operates with the same number of types as Holmarsdottir, but uses different names. Patton (2002: 80) argues that “...distinctions between paradigmatic, strategic and theoretical dimensions within any particular approach are both arguable and somewhat arbitrary” and uses the concept *Theoretic Traditions and Orientations*. He mentions 16 different research types. In addition he gives us an overview showing a number of scholars who operate with different ways of categorising.

I must admit that this made me a bit confused. It gave me, however, an understanding of the lack of consensus about how to classify, and the possibilities there are for the pragmatic and creative bricoleur to design the research according to her wishes.

Among the different possibilities I have chosen to use the concept *research types*, since I use the concept *research design* as a more complex concept (cf. 4.1), and I see the term *research approach* as a concept possible to use more generally. It is possible to classify my type of research as a case study. My case is the role and function of standardised testing.

#### *A case study*

When I decided to categorise my research as a case study, I did it because I have seen the concept being used in similar research. As an example Harvey (2005) calls his theoretical study of the Anglo-American influence on the global development of neoliberalism a case study. In the literature dealing with research methods, however, I found that cases very often were defined in a way that did not fit my case (Bryman 2004, Patton 2002). Stake (2005) opens up the definition which may cover my case. He argues that a “case may be simple or complex” at the same time it is “a specific One” (ibid.: 444). “In the social sciences and

---

<sup>52</sup> In the lecture 12.11.07, Holmarsdottir

<sup>53</sup> The concept is taken from Table 12.1: Characteristics of Four Qualitative Research Approaches in Johnson & Christensen (2000). *Evaluation Research: Qualitative, Quantitative and Mixed Approaches*. Boston, Allyn & Bacon. Copy of the table was received from Holmarsdottir during a lecture 12.11.07.

human services, most cases have working parts and purposes; many have a self. Functional or dysfunctional, rational or irrational, the case is a system” (ibid.: 444-445). He operates with three different kind of case study research: *intrinsic, instrumental and multiple case studies*, where he explains that instrumental case studies investigate some concrete examples in order to show trends. My concrete example is the PISA league tables. The production of and use of them exemplifies a trend; more concretely the growing use of standardised tests in the production of knowledge to be used in policy making within the field of education.

“[T]he primary reason for developing and conducting this large-scale international assessment [PISA] is to provide empirically grounded information which will inform policy decisions” (OECD1999: 7).

Stake (2005: 445) admits the following, which describes the general experience of lack of consensus within the field of qualitative research that opens up for the bricoleur: “Here and there, researchers will call anything they please a case study”.

#### *A critical case study*

As argued in chapter III, my theoretical angle has been critical theory, hence I have conducted a critical case study. My intention has been to question the growing test culture. This has been in line with Patton’s (2002: 131) following description: “[W]hat gives critical theory its name – what makes it critical – is that it seeks not just to study and understand society but rather to critique and change society”. Lather (in Denzin & Lincoln 2008: 5) puts it this way: “[C]ritical qualitative research represents inquiry done for explicit political, utopian purposes, a politics of liberation, a reflexive discourse constantly in search of open-ended, subversive, multivoiced epistemology”.

#### *Preparing for the field work*

Although my field was not geographically situated, I did conduct some field work; in the libraries, in the media, online, at conferences, in Union of Education Norway where I am a trade union representative, among people active in social movements and teachers in the classroom.

Before starting the field work I did some preparation:

- I divided my main research questions into more concrete questions, see appendix F
- I made an overview of academic journals of interest for my research.

- 
- I made a timetable for the study, see appendix G
  - I made an overview of *in depth documents*, see appendix H
  - I prepared my field work diary and an archive system where I could store and easily trace information of interest obtained from media, internet, conferences etc.
  - I read literature about how to code and analyse the data obtained from documents.

All this made the start of my study easier. During the process, however, the research questions changed slightly (cf. 1.2) due to deeper understanding of the topic. New literature was added, some texts were put aside. Also my plan on what to do when has been changed. As indicated in the timetable, I was supposed to finish my reading and analysis in January. But I never stopped to read and analyse. Reading and analysing have been a much more continuous process than expected.

#### 4.1.2 Data collection

In our modern society, much of social life is mediated by written texts that provide the qualitative researchers with material (Peräkylä 2005: 870). Moreover, documents are unobtrusive. These fact and the theoretical approach of my main objective, settled my decision to accomplish my research through documentary analysis. To be able to answer the research questions, I had to gather a variety of data from different academic fields as well as from current economic development. My sources were books, journals, and the World Wide Web. I also needed to follow the ongoing discussions in academia and the social movements, therefore I decided to use a grant given to me from Union of Education Norway, to attend as many conferences and seminars on issues related to my topic as possible, see appendix I. Moreover I followed the debates going on in the media and in Union of Education Norway, through their internal information and open accessible journals and web pages. I also attended the Union of Education meetings for shop stewards<sup>54</sup> in the Oslo schools, the meetings in the section for teachers in the compulsory school where I am a trade

---

<sup>54</sup> “A ‘shop steward’ would be a teachers' union representative in the school and responsible for taking the opinion or galvanising the opinion of other members of the same union in the school or calling them to (industrial) action. It's an expression more commonly heard in a factory or mine, but it is used in schools also” (Malcolm Mercer , Consultant in Education and Development, 07.05.2009).

union representative and the meetings in the committee of international affairs where I am a trade union representative.

### *Documents*

I divided the documents I used in two categories; the documents used as background information and the documents analysed thoroughly. The background information has been of general theoretical, political, economic, historical and pedagogical interest. The documents I analysed more thoroughly, in depth documents, have been used in the analysis of the background, construction and use of the PISA league tables, see appendix H. I tried to limit the number of in depth documents according to the advice of Bryman (2004: 393), who suggests not to sample more than between six and ten. My main sources directly connected to PISA have been the document that introduced the study: *Measuring Student knowledge and Skills, A New Framework for Assessment* (OECD 1999), and the latest PISA document available for me, *PISA 2006 Technical Report* (OECD 2009). Those two were consciously chosen because of my research focus and also because they for the time being frame the PISA study. Another main source has been the book *Pisa zufolge Pisa - Pisa According to Pisa* (Hopmann, Brinek & Retzl 2007).

Primary documents build on primary data collected or observed from a first hand experience, like the book of Bergesen (2006) and probably the raw data of PISA presented in OECD (2007a). Secondary documents rely on secondary data that are collected and processed by others than those with first hand experienced, like OECD (2007b), whereas examples of tertiary documents are indexes and abstracts<sup>55</sup>. Accordingly, I have used all the three different groups of documents.

Documents are divided into two types according to their intentions: Holmarsdottir used the concepts *solicited* while talking about documents made for a special use and *unsolicited* for documents produced for personal use. Duffy (1987: 54) makes the following division:

Deliberate sources, ( ) are produced for the attention of future researchers. Inadvertent sources, ( ) are used by researchers for some purpose other than that for which they are

---

<sup>55</sup> Holmarsdottir explained the differences between primary, secondary and tertiary documents, and the differences between solicited and unsolicited and restricted and unrestricted documents, in *Advanced course in qualitative research methods I*, the fall of 2008

---

originally intended. They are produced by the processes of local and central government and from everyday working of the education system.

I have used *inadvertent* and solicited sources, more specific official documents derived from state and private sources and mass media outputs, to use the categories of Bryman (2004: 386-390). I have also used archived research data and historical documents. Most of them have been *open*; freely accessible. The internal information from Union of Education Norway has been *restricted*, since it is information distributed to members and/or persons holding a position of trust. One of the documents that I intended to use as an in depth document, OECD (2007a) presenting the raw data of PISA 2006, ended up as a backdrop for my understanding of PISA. This is a *deliberate* source, since it is a general request to use it in more secondary analysis.

### *Interviews*

To support my analysis, I intended to interview unionists, headmasters and teachers about their experience with the growing number of standardised testing. When I had to narrow my focus to be able to manage within the frames of this thesis, I decided not to go through with interviewing. By that time I had accomplished unstructured interviews with two union leaders, Vice President in PGRI (Teachers Association of The Republic of Indonesia) Dr. Anna Sahaenah Soeparno and General Secretary of UNATU (Uganda National Teachers Union) Teopista Birungi Mayanja and three semi-structured and two unstructured interviews with headmasters from Oslo. I decided not to quote from them, but only use them as a general backdrop in the research process and save them for a possible later use.

### *Field notes*

I have been writing field notes regularly, mostly every day. They are a mixture of information gathered from documents and media and my own thoughts and reflections connected to formal as well as informal discussions in seminars and conferences, research documents, media notices and discussions, meetings in Union of Education Norway and peer discussions with students at the University of Oslo (UiO), teachers in the Oslo school and trade union representatives from Union of Education Norway and other unions. During the conferences and seminars I have had the possibility of engaging in formal as well as informal discussions

with respected scholars and activists. Michael Apple<sup>56</sup> as well as Yash Tandon<sup>57</sup> participated at Norway Social Forum, to mention two whom I have also quoted in my thesis.

For the last years, I have posed two questions to very many people: What is globalisation? What do you mean about PISA? Many of the answers from the last year are noted, but in the glare of belated wisdom, I should have started to make notes earlier. As it is now, I only can use the answers as a backdrop for my own thoughts.

Regularly I have read through the notes to keep all the different positions that are reflected there in mind. They have helped me to form questions and to search for literature as well as during my analysis. In fact I have experienced the field notes as an important data source during the research process.

### 4.1.3 Analysis

According to Kvale (1996: 184), “[t]o *analyze* means to separate something into parts of elements”, but there is no agreement within qualitative research on how this is being done. “Qualitative analysis transforms data into findings. No formula exists for that transformation. Guidance, yes. But no recipe. Direction can and will be offered, but the final destination remains unique for each inquirer, known only when – and if – arrived at” (Patton 2002: 432). However one “principle in document analysis is that everything should be questioned” (Duffy 1987: 57). Moreover it is our theoretical basis that guides our decision on how and what to analyse (Kvale 1996). Chapter III discussed my theoretical basis, and as shown there, I have chosen different theories to be able to shed light on my case and to reveal connections and build understanding through a holistic view on the issue<sup>58</sup>.

Patton (2003: 432) uses colourful metaphors while talking about what qualitative analysis is capable of doing. Metaphors that can make any researcher, especially a newcomer, humble - but at the same time they give a vision to reach for:

---

<sup>56</sup> Dr. Apple is a professor at University of Wisconsin, Madison, and the writer of many books <http://epaa.asu.edu/epaa/board/apple.html> (accessed 11.02.09).

<sup>57</sup> Cf. 3.3.1

<sup>58</sup> This is a free paraphrasing from a lecture conducted by Holmarsdottir 12.11.08.



Medieval alchemy aimed to transmute base metals into gold. Modern alchemy aims to transform raw data into knowledge, the coin of the information age. Rarity increase value. Fine qualitative analysis remains rare and difficult – and therefore valuable.

Metaphors abound. Analysis begins during a larval stage that, fully developed, metamorphoses from caterpillar-like beginnings into the splendour of the mature butterfly. (...) Findings emerge like an artistic mural created from collage-like pieces that make sense in new ways when seen and understood as part of a greater whole.

As a bricoleur I have tried to find the collage-like pieces by analysing the rhetoric used in the documents and seen it against an understanding of reality framed by theory, experience and a concept of common sense.

#### *Approach to the analysis: Hermeneutic*

Dialectics and critical theory are points of departure in this study. In search for an appropriate analysing approach, I found that the hermeneutics suited my intention best since;

- [it]is an approach to the analysis of texts that stresses how prior understandings and prejudices shape the interpretive process (Denzin & Lincoln 2005:27);
- [its]appeal to qualitative researcher is that it is an approach to the analysis of documents (...) that explicitly draws on two central tenets of the qualitative research strategy: an emphasis on the point of view of the author of the text and a sensitivity to context (Bryman 2004: 395), and
- [it fits into] Marxism, where the interpreter looks for meanings behind or beneath what is directly expressed (...) [and where] manifestations of an ideology concealing the basic contradictions of the social and economical forces at work (Kvale 1996: 203)".

#### *Discourse analysis*

Bryman (2004: 539) refers to *discourse analysis* (DA) as an “approach to the analysis of talk and other forms of discourse that emphasizes the way in which versions of reality are accomplished through language”. In his definition, one gets the understanding that DA is about oral language, but according to Peräkylä (2005), DA is also useful while analysing texts. Moreover, it also fits my contextual emphasis: “DA emphasizes the way versions of the world, of society, events and inner psychological worlds are produced in discourse” (Potter in Bryman 2004: 370) and “the discourses should be examined in relation to social structures, such as power relationships, that are responsible for the occasioning of those discourses” (Reed in Bryman 2004:377).

According to Peräkylä (2005), there are different forms of DA, two that fit into my analysis is *critical discourse analysis* (CDA) and *historical discourse analysis* (HAD). A researcher using CDA is “interested in the ways in which texts of different kinds reproduce power and

inequality in society” (ibid.: 871). My in depth documents are analysed according to CDA. I do, however, also place my documents into a historical perspective, where I look into the relationship between rhetoric of the documents and the reality where they function in connection to historical time and prevailing theories in society (ibid.: 872).

#### 4.1.4 Challenges

All my years as a teacher and also approximately 20 years of positions at all levels of the teachers union; school level, municipal / county level, national and even international level, make me an insider in the context of my study which pose challenges to my work as a researcher. To be an insider might be both an obstacle and an advantage. When my task is to analyse documents I choose to view it as an advantage that I have background knowledge from other sources than the written words. Nowadays few social scientists hold the view that a researcher could or should be without preconceptions (Bryman 2004). Patton (2002, p. 40) argues that “the researcher’s personal experiences and insights are an important part of the inquiry and critical understanding of the phenomenon”. Nobody is like a white sheet. Whether we wish it or not, there will always be a kind of subjectivity attached to what we do, also while doing academic work. However, it is important to be aware of this and to try to counteract our own biases.

#### 4.1.5 Validity and reliability or trustworthiness?

Important criteria to use when a quantitative research is evaluated, is whether it is valid, reliable and objective. There has been a long lasting debate on whether the same criteria should be used in qualitative research. According to Bryman (2004) and Kvale (1996) some researchers argue that the same criteria should be used, others ignore them. Guba and Lincoln (1985 in Bryman 2004) entered into this debate by suggesting that different criteria are needed in qualitative research by presenting an alternative terminology. They argue that *trustworthiness* and *authenticity* are criteria telling how good a qualitative study is. *Credibility*, *transferability*, *dependability* and *confirmability* are the four criteria making up trustworthiness. Brock-Utne (1996: 612) argues that “[t]he questions of validity and reliability within research are just as important within qualitative as within quantitative methods though they have to be treated differently”. I will not take a stand in this discussion.

---

*Internal validity/credibility*

*Credibility* parallels the concept *internal validity*. Where internal validity relates to causality, credibility asks whether such causality actually exists in social reality. In fact there might be multiple accounts of realities in the real life. To safeguard the credibility, it is important to make an effort to show an issue from different angles (Bryman 2004).

*Triangulation* is one way to attain credibility as well as internal validity. There are four basic types of triangulation; *data*, *investigator*, *theory* and *methodological triangulation* (Denzin 1978b in Patton 2002: 247). In this study I have used data triangulation, since I have gathered my information from different documents as well as through the process of being a part of the present discussion concerning the economic as well as the educational development. I have also used theory triangulation, since I have used different theories while analysing my data.

In Bryman (2004: 274-275) only the validation of respondents or members is mentioned. In my case, drawing data from documents and assessing them is one way of validating my findings.

While picking my documents I considered their authenticity, credibility and representativeness. To meet the demand for the latter, the in dept documents have been chosen to show two different comprehensions of PISA. To meet the demand for authenticity and credibility, the in depth documents are official documents and documents written by scholars, except the one written by Bergesen (2006). The background documents are mainly books written by scholars belonging to different academic subjects and articles from academic journals. A few books and articles are written by intellectuals and activists without an academic position. All the web pages used as references are official web pages operated by governments, organisations and scholars, except for Wikipedia that has been used in one footnote of minor importance about Alan Greenspan.

*External validity/transferability*

External validity tells us about to which degree findings are generaliseable across social settings (Bryman 2004). Transferability parallels external validity by asking whether the contextual in a qualitative study make it possible to generalise. Transferability or external validity may become obtained through thick descriptions that might make the knowledge that

is drawn from the findings in one milieu possible to “transfer” to another. This text is almost as “thick” as possible within the frame of this study.

According to Brock-Utne (1996) a component of external validity is ecological validity. The less anonymity, the higher is the ecological validity. I have mainly used documents accessible to everyone in my study, thus the ecological validity should be high.

### *Reliability/dependability*

Reliability occurs when measures used in social science are consistent and a study is repeatable (Bryman 2004). Dependability parallels this concept and questions if it really is possible to develop consistent measures independent of changing contexts in social settings. “The commonly held assumption that qualitative methods pay attention to validity and not to reliability is false” Brock-Utne (1996: 612) argues. “What is important is the fallibility of any method” (ibid.).

To attain dependability, Bryman (2004) suggests that the researcher should act as auditors, keeping records through out the research for other researchers to be able to repeat the study. As mentioned, I made an archive and have been writing field notes. The latter is too personal to be of any use, the archive I will keep. Moreover, in a documentary analysis, the reference list is also a type of archive.

### *Objectivity /confirmability*

Confirmability parallels the concept objectivity, and argues that complete objectivity is impossible to attain in social science. My position is that objectivity is an illusion. I do agree with Steven Klees when he argued that “Research is biased”<sup>59</sup>.

What is possible for the researcher to show, however, is that she has been acting in good faith (ibid.). My answer to the demand for confirmability has thus been to lay the cards on the table, as Kvale (1996) puts it. I am aware of what I bring into the study and how that might put a bias to it. Yet I have tried to be open to different view-points, read not only critics of PISA but also those who have been in the forefront defending PISA. But by showing my own theoretical framework clearly in line with critical theory, I hope to have

---

<sup>59</sup> In the discussion following the seminary *Global monitoring, International Aid, and Country Development strategies*, CIES 2009, 24.03.09.

---

been able to show how certain perspectives will follow from this framework. Thus studies relying on critical theory

legitimate an important approach to theoretical or ideological elaboration, confirmation, and elucidation. What is required is that the researcher be very clear about the theoretical framework being used and the implications of that perspective on study focus, data collection, fieldwork, and analysis (Patton 2002:191).

## 4.2 Summing up

In this chapter as well as in chapter III, I have been laying my cards on the table. In the chapters to come I will use them, as well as the pieces gathered from my data collection, to make a bricolage. Chapter V, *the Wheel*, will try to draw parallels between the economic and the education development from WWII and up till today. Chapter VI, *the Hub*, will use PISA as an example while criticising the epistemological, ontological and methodological positions dominating the current knowledge production. The issue is to find out if standardised test can be seen as *the Hub in the Wheel*.

## 5. The “wheel”

“[E]ducation *is* political” Brock-Utne (2006: XVII) argues, and the political struggles about ideology, goals and methods occur at all levels in all educational settings, education systems and communities. Different stakeholders with different interests, such as students, union leaders governmental or business representatives, to mention some, enter into this struggle. “These struggles not only shape educational policy and practice, they are also dialectically related to more general relations of power along social classes, racial/ethnic groups, gender groups, and nations” (ibid.). It is the outcome of those struggles that shapes the educational systems where “existing social relations are reproduced, legitimated, challenged or transformed” (ibid.).

In chapter III I sketched the process of neoliberalisation; how it was anchored in the creation of think tanks, reinforced by the displacement of power after the breaking down of the USSR and supported by the big post WWII organisations. Chapter II gave an overview of the development of a test culture within education. My research question indicates a possible connection between the neoliberalisation project and the development of this culture. In this chapter I will try to uncover this connection by drawing some parallels between politics, economy and education.

### 5.1 *Left vs. Right*

Seen from a dialectical approach, the context is essential. As shown in chapter III, the economic context of educational development in the western part of the world has varied in the post war period. Immediately after WWII embedded liberalism, curbed capitalist power, economic growth, redistribution of wealth and the creation of the welfare states was the backcloth of the development, while for the last quarter of the century it has been neoliberalism (Harvey 2005). How did this contextual change influence the educational development?

### 5.1.1 The period of embedded liberalism

I have chosen the heading *left* vs. *right* for this section; a dichotomy to underline the ideological span existing in the *post war* period, although it did not remain completely dichotomized in the real world. This ideological difference that existed between the US and the USSR can be seen as a reason for the tense political power struggle between them, which led to geopolitical rivalry and tragic wars. The two superpowers never entered into direct military confrontations, thus the period is also called *cold war*. Internationally the interpretation of this term might seem rather arrogant, since lots of countries were thrown into most real wars at that time, such as Korea and Vietnam to mention two. The situation in the cold war period polarized the way many people thought, it forced people to take a stand; the *west* represented the *right*, the *east* the *left*. At the same time the power relations between the US and the USSR and the western fear of the *other*, created the more human capitalism during the embedded liberalism. Wahl (2009) builds his analysis of the *post war* development of class compromises and welfare states on the power relations in that period.

This ideological, political and economic context had a trickle down effect on different fields and marked a period of quite visible frontiers also within educational theory and philosophy, particular in the 1970s (Husén & Tuijman 1994: 4, Lawn 2008b). Worldwide, the *post war* period was characterized by the *south*'s fight for liberation from colonialism. Traditionally, the colonies had followed the school and evaluation systems of their masters (Cummings 2003: 205), but now voices emerged from the *south* questioning western schooling and evaluation systems. One important voice was Paulo Freire (1985, 1993) from Brazil. Rikowski (2007) uses 1970, the publishing year of the English version of his book *Pedagogy of the Oppressed*, as the starting point of what he calls the period of *the Old Marxist Education Theory*<sup>60</sup>. The book of Freire is seen as essential for a *left* influence within educational theory internationally (Jones *et.al.* 2008, Rikowski 2007, Small 2005). Another voice from the *south* was the president of Tanzania, Julius K. Nyerere, who in the essay *Education for Self-Reliance* (1968) presented a radical, holistic educational policy for his country. The rediscovering of the Bolshevik educators from the 1920s (Jones *et.al.* 2008: 9),

---

<sup>60</sup> According to him this period ended in 1982 due to the Marxist scholars own lack of capability to address basic educational questions.

among others Lev Semenovich Vygotsky, also contributed to an alternative non western view on education.

“[S]chools in Western Europe were for a period home to values and practices embodying solidarities of a sort resistant to the logic of the market” (Jones *et.al.* 2008:2). They further argue that the development was driven by state-focused, welfare-oriented movements trying to create an educational system based on increased equality and social citizenship. The 1968 generation teachers “came to think of school as an institution where democracy, cultural recognition and equal opportunity could serve as central principles” (Ibid: 9). In Norway the same ideas at the same period of time fostered an alternative and anti authoritarian university education called social pedagogy, where importance was attached to action research, students’ involvement, cooperation, empowerment and the development of critical thinking (Brock-Utne 2008b).

As mentioned in chapter II, the 1970s was a period when a more holistic view on education was discussed and to a certain extent accepted. Although the development of all sorts of tools usable for evaluations and monitoring had been developed, the scepticism towards the use of them was growing also in countries where they earlier had been accepted (Husén & Tuijman 1994). From a theoretical perspective these were the years when the Marxist theory was visible. From an epistemological viewpoint, those were anti positivistic years and from an ontological viewpoint they were dominated by constructivism. Seen from a qualification perspective, there was a wish to stress creative qualifications and to diminish the development of the adaptability qualifications. In short, those were the years of a radicalisation of educational ideas and to some extent practice; the years of the so called progressive pedagogy was developed. However this was not a long lasting period.

### 5.1.2 The period of neoliberalism

After the new direction in China’s policy and the fall of USSR as mentioned in chapter III, the power balance between the *east* and *west*, or between the *left* and *right*, were changed. The regulations that were built to curb capitalism fell one after another, and soon we were on our way towards neoliberal capitalism. This happened not only in the *west*, but internationally (Harvey 2005). The development was presented by Thatcher as unavoidable; *There Is No Alternative*. I will argue that the new economic policy also changed the direction



---

of the educational policy, and later in this chapter and in the chapter to come, I will try to show the relevance of the following allegations: From a theoretical perspective the postmodernist and poststructuralist theories building on structural- functionalism became more visible, from an epistemological viewpoint, positivism in the shape of neo positivism appeared and the ontological position turned in an objectivistic direction. Seen from a qualification perspective, creative qualifications became qualifications for the few, and emphasise on adaptability qualifications reappeared.

### *Post neoliberalism?*

After the financial crisis of 2008 has turned into a more general economic crisis, everything might happen. As a matter of fact the contours of something new within the economic development are visible. Only time will tell how this new context will influence the education development in the future.

### 5.1.3 No differences between the *left* and *right*?

While the period of embedded liberalism was characterized by ideological and political polarization, the next period was blurred. The educational transformations during the neo-liberal offensive were “brought about by joint work of governments from the left and right” (Jones *et.al.* 2008: 20). Prime Minister Tony Blair from Labour continued the politics of John Mayor and Margaret Thatcher from the Conservative Party (*ibid.*), and I will argue that the Norwegian ministers of Education from the *left*; Øystein Djupedal and Vegard Solhjell, mainly continued the politics of their predecessor from the *right*; Kristin Clemet<sup>61</sup>. In 2005, the red-green coalition came to power, and two years later they launched a new white paper on education *...and no one was left behind*<sup>62</sup> (St.meld. nr.16 2006-2007<sup>63</sup>). Whereas the essential news in this strategic plan is to emphasise on early intervention, a thought borrowed from Finland after their successful PISA tests, the way to uncover the need for intervention is

---

<sup>61</sup> More about her role will come later in this chapter.

<sup>62</sup> In Norwegian: *...og ingen sto igjen*, a title almost similar to the governmental document of George Busch; *No child left behind*.

<sup>63</sup> The abbreviation used when referring to a white paper is *St.meld.* in Norwegian. From now on I will use the Norwegian abbreviation when I mention a Norwegian white paper.

through centralised mapping tests. In addition the National tests were reintroduced, although in a limited number and in different standards<sup>64</sup>.

The latest white paper; *Quality in school*<sup>65</sup> (St.meld. nr.31 2007-2008), follows in the same direction; more mapping tests were introduced. When the white paper was approved in the Norwegian Parliament on the 21.11.08, Gunnar Gundersen from the right wing party was pleased because according to him, it put an end to the educational policy of the 1970s (Nilsen 2008: 8). He was supported by Ine Marie Eriksen Søreide, also from the Conservative party and the leader of the executive education committee in the Parliament, arguing that the red-green coalition carry out an educational policy of the right (Bredeveien 2009: 7).

Some argue that the blurred positions are due to the influence of the post modernists and their reluctance to take a stand. “True” answers do not exist. Everything floats. What counts is the present; a present that has become so difficult to understand that to take a stand is impossible (Marsdal & Wold 2004). They also proclaim postmodernism to be the “lubrication” of neoliberalism; without this ideological position, the changes introduced by neoliberalism would have been resisted by more people, and hence not so easy to carry through. Rikowski (2008) argues that the postmodernists’ resistance to make up their mind leads to a dereliction in a time where explanations are much needed.

## 5.2 To change the way people think

In chapter III I used Harvey’s (2005) analysis of the importance of neoliberal think tanks and the export of neoliberal economic ideas in the process of changing the way people thought. I chose the analysis of Harvey because I found it interesting in three ways. At first I could not avoid thinking about the *Inquiry – IEI* - from chapter II while reading about the think tanks. Secondly I found some similarities between the hegemonic role economic liberalism was playing and the growing belief in the test culture within education. Thirdly I simply thought about the potential power of research and the potential use or misuse of research results.

---

<sup>64</sup> See chapter II and appendix E

<sup>65</sup> In Norwegian: *Kvalitet i skolen*

### 5.2.1 Think tanks

IEI is seen as the first “modern” scientific networking in education, gathering an elite group of approximately 100 participants discussing research theories and methodologies. They were psychologists, progressive educators and comparative educationalists holding different perspective. The more holistic perspective of the two latter groups was however during the project “overtaken by the creation of measurement standards, usable in comparing schools, regions and even countries” (Lawn 2008a: 23). Lawn indicates some answers to why the development went in such a direction in his article. He argues that philanthropic foundations have been important in educational research in the US, and also in countries within the sphere of the US without corresponding foundations, like New Zealand, South Africa and Australia. Husén & Tuijnman (1994) mention the role of Ford foundation and Lawn (2008a: 20) mentions Carnegie as an example of a foundation that extended “its field of action in education in the 1920s, into cross-border actions in Europe, the British empire and even China”. The extensive funding of IEI can serve as an example of the European involvement. Was the contribution of Carnegie to educational research an altruistic action?

Lawn (2008a) uses research done by Lagemann (1989) on the Carnegie Cooperation while trying to view their contribution in a broader perspective. The foundation’s approach seemed to be that experts independent of governmental control were the best “to deliver fair and objective knowledge” (Lawn 2008a: 19). At the same time they had a wish to safeguard the US’s “traditions of liberty and individualism” (Lagemann in Lawn 2008 a: 19) and to protect “Anglo-Saxon ‘race’ privileges” (Lawn 2008a: 19). A neighbouring question is whether the “objective” knowledge should serve their subjective wishes? Although underlining all the uncertainties around the project, Lawn indicates that Carnegie picked American experts sharing their thoughts, and that the rest of the participants in addition to their knowledge, were picked because of common circles of acquaintance.

At the time of the *Inquiry*, the US had replaced the earlier position of Germany, and had become dominant within “the cosmopolitanism of an elite group of educational researchers (...) and the ‘internationalisation’ of research theories and methodologies”. Some of the key American actors in IEI had “a clear sense of mission to modernise examining practices in other countries (in the American way)” (both quotations in Lawn 2008a: 20). Strien (1997 in *ibid.*: 23) sees it as a sort of “scientific colonisation”, an “intellectual domination of an existing culture by a foreign, more powerful culture”.

In the vocabulary of today, IEI might be seen as an American/European think tank. Although the ideas came from the US, they had a sounding board within the European academia, especially within the UK, and with the money from Carnegie they could be further developed. In chapter II I covered the influence of IEI in the post war period. From a 2009 perspective, we see that what the inquiry laid the foundations to are elements of educational thinking nowadays; the scientific Esperanto, the dominating role of English – and in addition the feeling that *There Is No Alternative*. However the most burning question is whether this system gives privileges to some by protecting the race, traditions of liberty and individualism of the Anglo-Saxon world as Carnegie wanted -? Is this also a system widening the gap as we have seen in the case of neoliberalism in chapter III?

### 5.2.2 The use of test results

Comparisons between countries have been the cause of change in national policy several times. In chapter II I mentioned the Sputnik shock, where the USSR ability to launch a rocket into orbit caused huge furore and subsequent fundamental change in the education policy of the US and also lead to the demand for the development of international comparative studies (Husén & Tuijnman 1994:5). The next crisis happened in the US when the results of SIMS<sup>66</sup> were published. Although the study was much broader, the media focused mainly on the low achievements in the US and some West European countries shown by “the comparisons of the mean performance of 13-year-old students in arithmetic, algebra, and geometry” (ibid.:10). This and the main concerns in the report *A Nation at Risk* “served to focus the attention of influential stakeholders on the use of measures of student achievement as criteria for judging the performance of the education system” (ibid.: 11).

The *A Nation at Risk* report played a pivotal role in raising concern over the quality of education. In a more subtle way it also raised political awareness that the international comparative studies of student achievement could be used for the benchmarking of performance levels in different education systems (ibid.: 10).

The US are not the only country that has experienced how the use of the results of comparative studies has influenced the development. Sjøberg (2006:195) mentions that the

---

<sup>66</sup> SIMS , see appendix B.

---

World Bank uses studies tailored to the same pattern as TIMSS as a precondition for support to the education sector in countries in the *south*. Dolin (2007) shows the Danish reactions to the media headlines after the PISA survey of 2003, Sjøberg (2006, 2007) and Bergesen (2006) do the same seen from a Norwegian perspective and Sjøberg (2007: 208) shows us war like headlines from a German newspaper in 2000. In the following I will look into the Norwegian experience.

### 5.2.3 The use of test results in Norway

In the 1970s there was a widespread scepticism towards the positivistic belief that everything in life is possible to measure, weigh and count. Pedagogues critical to positivistic solutions, argued that it was an expression of a behaviouristic and instrumental way of thinking. As we have seen in chapter II, this was also a position taken by Parliament at that time. However, in 1982, when we had a right wing government, Norway registered for SISS<sup>67</sup>, and in the 1990s, during a social democratic government, Norway registered for more tests. TIMSS and PISA were two of them. The Norwegians were moving from a sceptical position towards statistics to a position valuing statistics as the mirror of the “truth”; “from number phobia to the magic of numbers”, as Sjøberg (2006: 190)<sup>68</sup> puts it.

In the following I will concentrate on the PISA results, since as Sjøberg (2009: 26) argues; “all educationalists know that PISA has delivered the premises for the debate in the [Norwegian] media and is also the study most often used by the politicians”.

PISA is an international study that grows bigger each time it is administrated. It is also a broad study, looking into different aspects of schooling. One “product”, coming out of the study is the tables that compare the results from different countries. Hopmann, Brinek & Retzl (2007) name them league tables, a term I also will use. Just as in the case of the US, the media emphasise on those tables every time the PISA results are published. In chapter II we saw that PISA assesses reading, mathematical and scientific literacy and that it has different foci every three years. In 2000 the focus was reading literacy. Although the Norwegian researchers responsible for the studies in Norway were cautious not to make

---

<sup>67</sup> SISS, see appendix B.

<sup>68</sup> My translation of the expressive heading in Norwegian: “Fra tallfobi til tallmagi”.

generalised conclusions out of the league tables, like Kjærnsli and Lie (2006), the tables were used by the media to draw a disastrous picture of the Norwegian school performance and system as a whole. In the league table showing reading performances<sup>69</sup>, Norway ended as a mean performing OECD country; number thirteen. However, only seven countries performed significantly better than us (Roe 2006: 179). Nevertheless, massive attacks on the Norwegian school system mainly based on this and similar league tables, shaped the consciousness of the public and made it easy for the right wing Minister of Education Kristin Clemet, to start the process of reshaping the Norwegian education system. In his book, the Parliamentary Secretary of the Minister of Education, Helge Ole Bergesen (2006: 12), also from the Conservative Party of Norway expresses gratitude to the journalists, both those who understood what happened - and all the others - for keeping the heat of the debate. He also describes 4<sup>th</sup> of December 2001 under the heading “The PISA shock”; when the researchers Astrid Roe and Svein Lie not explicitly but with their laconic comments increased the appetite of the journalists, and when Clemet made their day with the following comment: “This is disappointing, almost like coming home from a winter Olympic game without a single medal. And this time we can not blame the others for using drugs” (Bergesen 2006: 41). The next day the forth power of the state did their job.

The new white paper on education, *Culture for Learning* (St. meld. nr.30 2003-2004)<sup>70</sup> was followed by the curriculum plan named *Knowledge Promotion*<sup>71</sup> (Kunnskapsdepartementet 2006). This represented a fundamental change of perspective regarding the concept of knowledge, where emphasise on basic knowledge was replaced with emphasise on basic skills (Østerud 2006). Sjøberg (2006) notes that the changes were triggered by a too generalized and too simplified use of statistics. Østerud (2006: 216) characterises PISA- and TIMSS as political documents. One may add; political documents that have become actors in their own right.

Union of Education Norway carried out an inquiry on Norwegian teachers and headmasters view on PISA the autumn of 2008. The result was published in the professional paper *Bedre*

---

<sup>69</sup> <http://www.oecd.org/dataoecd/31/19/34107978.pdf> : 5 (accessed 07.03.09).

<sup>70</sup> In Norwegian: *Kultur for læring*

<sup>71</sup> In Norwegian: *Kunnskapsløftet*

*skole*<sup>72</sup> in March 2009. Hjetland (2009) comments on the role of the media in the PISA debate, and argues that it is strange and disappointing that neither researchers nor school authorities enters the debate to make the picture made by the media more nuanced. Sjøberg (2009: 26, 27) has also commented the result of the inquiry, and his summary of the answers regarding the use of the PISA results is as follows:

- 81% are up in arm about how the media and the politicians misunderstand and distort the PISA results.
- 84% agree upon that politicians and school authorities pay to much attention to PISA results.
- 93 % agree upon that some politicians are extremely selective in their reading of international studies and forget data that gives other and more positive picture of the school.
- 77% agree upon that media fail to present studies that give a positive picture of the school.
- 83% agree upon that Norwegian school authorities often are uncritical and excepting when it comes to international studies and advises from foreign experts.
- 82% agree upon that the PISA debate gives a wrong picture of the quality of Norwegian schools.
- 54% doubt if it is the curriculum planes or the PISA results that guide the school.
- However 95% also agree upon that Norway to a greater extent ought to use research and knowledge while developing the education policy, and 50 % agree upon that Norway should use the results of international studies when formulating education policy.

The last ballpoint shows that the Norwegian teachers and headmasters are not negative to research as such, but that they are sceptical to the use of PISA results.

The same inquiry is commented by Kjærnsli and Roe (2009) who dismiss the whole inquiry due to low answering percentage, imprecise questions and interpretations. Sjøberg (2009) does not have any comments in the same directions. The two different position taken by the *insiders*; Kjærnsli and Roe, and the *outsider*; Sjøberg, is interesting but not unique, thus the positions in favour of or against PISA are indeed polarised; from praising to blaming. A problem discussed by Hopmann & Brinek (2007) is how the scholars working with the study acts while meeting criticism. The behaviour is not unlike the one that occurs in “large companies when they encounter a potential scandal, e.g. pharmaceutical companies dealing with ill-conceived drugs (ibid: 14)”. They argue that this behaviour has been seen in many European countries. (Hopmannn, Brinek & Retzl 2007).

---

<sup>72</sup> In English: *Better school*. Published by Union of Education Norway.

### 5.2.4 From Homo Politicus to Homo Economicus

In a neoliberal universe, human beings are not political creatures, but rather looked at as consuming, competitive and selfish individualists. This view is supported by media, advertising, political and economic analyses and theories, and has gradually become the way in which people look at themselves (Bank 2006: 242). At least many *uppers* in the western hemisphere, I would add. Moreover, the society is not built on solidarity and common values, but on self-interest. *There is no such thing as society*<sup>73</sup> is one of the well known utterances of Margaret Thatcher. One of the cornerstones is the understanding of human nature as competitive. Such an understanding will obviously influence a person's view on education. Competitive situations must be created on all levels to urge students, teachers, administrators and – nations - to do their best. As we will see later in this thesis, this is not an alien idea in the present situation. However, is competitiveness a part of human nature? Does competition motivate us to yield our best – and does it ennoble the man by building character? In his book *No Contest*, Alfie Kohn (1992) uses almost 250 pages to negate those questions.

In short he argues that the research done on competition as a part of human nature, takes the myths that the assertion is build on as a point of departure, and thus does not question them. Research on animal behaviour has shown that animals survive because of cooperation, not because of competition. “Natural selection does not require competition; on the contrary, it discourages it. Survival generally demands that individuals work with each other rather than against each other – and this includes others of the same species as well as those from different species” (Kohn 1992: 21). Researches on competition vs. cooperation among human beings show clearly the merits of the latter. Competition is central to Western culture and it is learned. Only the winners gain by this system.

It might be time to question one of the cornerstones on which we have built our conception of human nature and thus the present development, unless we are willing to continue to build it on a life-lie!

---

<sup>73</sup> The famous saying of Prime Minister Margaret Thatcher from Women's Own Magazine October 31 1987: <http://briandeer.com/social/thatcher-society.htm> (accessed 12.05.09).



---

*The new language*

A whole range of concepts have gradually entered into our language. According to Michael Apple<sup>74</sup>, they have contributed to a change in our common way of thinking; our common sense. I have already mentioned the concept “globalisation”, that “has come from nowhere to be almost everywhere” Giddens (1999 in Crossley and Watson 2006: 53). This concept might be regarded as *created* during the process of neoliberalisation.

Other concepts have been *co-opted* during the same period. Apple<sup>75</sup> argued that the understanding of some concepts, like *choice* and *knowledge*, has been co-opted by neoliberals; “emptied” and then refilled with new understanding. Those basically positive concepts have become neoliberal keywords; choice has even become *the* mantra. The facilitation of this individual freedom to choose has impact on the whole school structure. To be able to choose among schools they have to be evaluated and monitored. The results have to be published, and the money has to follow the students. This applies to systems where schools mostly are public like in Norway as well as in systems where public and private schools are competing, like in the US. In some parts of the world they have developed a voucher system like in Chile (Valverde 2004) and some states in the US (Klees 2008), other places a system of capitation grant facilitate the free choice, like in New Zealand<sup>76</sup>

*Knowledge* has become another mantra. Nobody can argue that knowledge is not important, however what is most often forgot when politicians argue and media repeat, is that knowledge is a multiple concept. The demand to measure knowledge has contributed to the narrowing of the concept. What kind of knowledge and whose knowledge counts are closely connected but hardly discussed. I will come back to this in chapter VI.

Many more concepts have been co-opted, included in the new rhetoric and placed “at the heart of the neo-liberal programme”, as Jones *et.al.* (2008: 138) argue. One of them is *decentralisation*; a concept that originally was used to counter bureaucratic and fordist hierarchies. The “old” understanding of the concept was guided by democratic way of

---

<sup>74</sup> Apple at Norway Social Forum 09.11.08.

<sup>75</sup> Apple at Norway Social Forum 09.11.08.

<sup>76</sup> Lecture 01.04.08 by Jon Lauglo

thinking – more power to the people – the “new” is to provide a competitive environment (Jones *et.al.* 2008).

Other concepts that have been co-opted are, according to Jones *et.al.* (2008), *autonomy*, *civil society*, *inclusion* and *creativity*. They have all been translated into free market terms, but can be traced back to nearly opposite origins. Autonomy was as an example used by liberals and socialists for “the independence of schools from outside interests such as the church or economic factors” (ibid.:138).

Other concepts are *imported* from the business sector, like those we have received through NPM and NPM like models. To dive into the creation of the new language; the imported, the co-opted and the newly created concepts, could actually have been an interesting topic in itself, but within the frames of this text I have to limit the discussion. Nevertheless a concept like *accountability*, which I explained briefly in chapter I, is worth mentioning. Husén & Tuijnman (1994: 4) underlines the connection between monitoring and accountability, saying that monitoring “is an important element of evaluation in a model of accountability”. Kohn (2000) argues that it is a rather new word in educational discourse and that it builds on the belief that Homo Economicus needs carrot and stick to perform; in the case of accountability, a stick. It enters into a cause and effect relationship as both a means and a goal, and explains the whole need for schools to bureaucratised. If people believe we need accountability in the neoliberal understanding of the word, many questions will remain unasked.

### 5.3 To change what people do

Neoliberalism is also called market liberalism because of its emphasis on the market as both a means and a goal. This market oriented way of thinking has also influenced education, and the process can be called the marketisation of education. Steps have been taken to adjust the education sector to a more neoliberal reality. One of them has been the creation of markets - or quasi-markets – where they earlier did not exist. In this section I will look into vital organisations and structures that have been active in the process of marketisation of education as well as in the process of the creation of markets. I will also look into what kind of qualifications that are produced in the age of neoliberalism.

### 5.3.1 Vital organisations

A number of powerful supranational organisations were established after WW II (cf. 3.1.3.). They have all been important allies in the neoliberalisation process. Some have also managed to remark schooling with their policy orthodoxy, as Jones *et.al.* (2008) formulate it in the case of Western Europe. I will add in the entire world.

#### *WTO/GATS*

I will start with WTO, since this is an organisation playing a superior international role with a very clear neoliberal bearing. In the words of Brock-Utne (2008b: 91), “the World Trade Organization (WTO) has become a vehicle for assuring that practically the whole world is open for the unhindered operations of private capital”.

WTO regulates international trade. The idea of making international trade rules for agricultural or industrial commodities is understandable, since those sectors of economic life produce concrete products simple to sell and buy in a market. However, one of the WTO agreements is *General Agreement on Trade in Services* (GATS); a treaty that has the ambition to regulate trade in services worldwide (Bank 2006). The service sector includes everything from health care to hair cut – including education. How is it possible to make international rules that regulate such services? They are not all commodities as such, are they? And what about the markets where the trade is supposed to happen, where are they? The former General Secretary of WTO Renato Ruggiero (1998) might have a point when he argues that GATS “affects areas never earlier regarded as trade policy. I think neither governments nor industries have understood the full consequences [of this treaty]” (in Seierstad 2005)<sup>77</sup>.

Education, and also indirectly education related services like library, culture, text book production and copyright etc., are included in the services GATS intend to regulate (Bank 2006: 236). To make services trade able, two premises have to be in place; the services must be defined – or redefined – as commodities, and markets where to sell the commodities have to exist or be created. One may say that this was done during the GATS negotiations in the late 1980s and early 1990s, also called the *Uruguay Round*.

---

<sup>77</sup> From a speech held 02.07.1998.

<b>Table 5.1. The Uruguay agreement</b>	
<b>The Uruguay agreement<sup>78</sup></b>	<b>The Uruguay agreement as seen in education<sup>79</sup></b>
For the purposes of this Agreement, trade in services is defined as the supply of a service:	The agreement unfolds when:
a) from the territory of one Member into the territory of any other Member;	an institution establishes an educational programme or institution in another country, for example through long distant learning
b) in the territory of one Member to the service consumer of any other Member;	students are getting their education abroad
c) by a service supplier of one Member, through commercial presence in the territory of any other Member;	one country establish an institution inside another country
d) by a service supplier of one Member, through presence of natural persons of a member in the territory of any other Member.	an educational institution have employees or take on a task in a foreign country.

All levels of education; primary, secondary, tertiary, adult and other (Bank 2006: 238), might enter into one or more of the main areas. When the agreement was sign in 1994, the member countries themselves decided what levels they wanted to open for competition at the international arena. Among the 30 countries in the world that opened up their education sector, Norway, at that time lead by the Social Democrats, Tsjeckia, Slovakia, Sierra Leone and Lesotho opened up the most. The result is that Norway has undertaken commitments in GATS at all levels, including compulsory primary education as well as tertiary education (Hjetland 2004).

A liberalised education system has to follow some main principles (Bank 2006). First it is the principle of *comparative advantages*. There are some nature-given aspects that make production of a commodity more effective in one country than another, and thus make the country more competitive. In the case of education, the Anglo- Saxon world has a

<sup>78</sup> The left column is a quotation from Article I in the GATS agreement: [http://www.wto.org/english/docs\\_e/legal\\_e/26-gats\\_01\\_e.htm](http://www.wto.org/english/docs_e/legal_e/26-gats_01_e.htm) (accessed 04.03.09).

<sup>79</sup> Bank (2006: 237) has made the explanation of how this article might apply for the field of education.

comparative advantage because of their language. This advantage is of special interest when it comes to student over sea business (Yang 2003).

Another principle is that of *non-discrimination*; all economic actors have the right to similar treatment. In education this will comprise the right of foreign educational concerns to establish and demand equal treatment when it comes to governmental support. To give preferential treatment to national institutions to secure a more equal society, as the Norwegian educational system has tried to do, might become illegal.

A third principle is the one attending to *diminish trade hindrances*. In Norway we have decided that it is illegal to earn profit on compulsory education. Could this position be questioned as a trade hindrance?

The finale principle is *continuous liberalisation*. When Norway signed the GATS treaty in 1994, we made some reservations connected to the former points. Will it be possible to keep those reservations when we at the same time agreed upon the principle of continuous liberation?

Due to huge disagreements and opposition<sup>80</sup>, the next WTO/GATS negotiation, the Doha round has not yet been signed, although they have been going on since 2001<sup>81</sup>. How influential GATS will be in the future therefore remain to be seen. However, I found it important to show the undressed neoliberal way of thinking at education as a commodity. If the struggle against the powers that wish for such a development is not continued, the fundamental ideas of GATS, those stating that everything should be looked at as commodities possible to trade in a market, might become normative in the future.

### *The World Bank/IMF*

The economic policy of the World Bank and the IMF have had impact on governmental spending and influenced the public sector where education is a part mainly in the south and in the former east block countries in Europe:

---

<sup>80</sup> Educational International (EI) is one of the organisations that have been fighting GATS since the end of 1990s, according to Dr Antoni Verger I Plannells from University of Amsterdam a quite successful fight, showing that is possible to “throw sand into the machinery” (From the meeting *Comparative Studies in Globalization and Education*, CIES 2009, 24.03.09).

<sup>81</sup> [http://www.wto.org/english/tratop\\_e/dda\\_e/dohaexplained\\_e.htm](http://www.wto.org/english/tratop_e/dda_e/dohaexplained_e.htm) (accessed 21.04.09).

Directed by Western interest the World Bank and the International Monetary Fund have used and continue to use their creditor powers to pressure first the poor debtor countries in the South and then the collapsing members of the former Soviet Union to turn their own battered economies into the same kind of unrestricted markets (Brock-Utne 2008 b: 91).

The one size fits all advices given by the organisations have been heavily criticised (Breidlid 2006, Brock-Utne 2006, Samoff 1996), as well as their political and ideological position. I let Klees (2008: 312) represent the critique: “World Bank policy (...) is the mayor player in global educational policy and has been at the forefront of the shift to neoliberal thinking”.

The idea of the existence of a budget cap has been central in neoliberal thinking. This idea has lead to the downsizing of governmental spending and further to a competition between different public sectors; a “beggar-thy-neighbour policy” as Klees (2008: 318) calls it.

Moreover it has lead to the cut in real terms of expenditures per head in approximately two-third of the countries in SSA. In Tanzania the cut was 60%. The lack of money allocated to education led to more private schools, introduction of school fees both in public and private schools and decline of enrolment, especially for girls (Brock-Utne 2006). One of the World Bank/IMF advices was based on a cost benefit analysis that showed highest *rate of return* (ROR) when investing in primary education (Klees 2008). As a result less governmental and aid money was allocated to higher education, tuition fees were introduced and private higher education institutions were promoted. In 1986, the World Bank even argued that higher education in Africa was a luxury (Brock-Utne 2006). In other words, Africa does not need to produce intellectuals! Brock-Utne (2006: 46) is not prudent in her critique:

[T]he main goals of the World Bank’s efforts in Africa, both in terms of sectoral lending and as components of structural adjustment programs, seem to be to prepare people for the jobs that a global division of labour offers, which means primarily producers of raw material.

Seen in a WTO perspective, the production of raw materials might be the comparative advantage in SSA. Seen from the perspective of the economist Reinert (2004), this way of thinking will lead to regional stagnation.

Many scholars argue that the Bank and the IMF have done huge damage to the educational sector in the *south*, such as Brock-Utne (2006, 2007b, 2008b), Klees (2008), Mazrui (1997), Samoff (1996), Yang (2003), to mention some. However after the damage is done, Klees (2008) argues, the World Bank and the IMF by now recognize that taxation is needed to build a sustainable education system and that the ROR analysis was erroneous.

---

## OECD/EU

Nowadays the 30 wealthiest countries in the world are members in the organisation that openly confess to promote the market economy:

OECD brings together the governments of countries committed to democracy and the market economy from around the world to:

- Support sustainable economic growth
- Boost employment
- Raise living standards
- Maintain financial stability
- Assist other countries' economic development
- Contribute to growth in world trade.<sup>82</sup>

Simultaneously, the OECD has a long history of engagement in education. As seen in chapter II, the organisation was in the forefront of the development towards more monitoring and international comparisons. As early in 1973 they formulated their first educational indicators, although the time was not ripe for the implementation of them yet. Since 2000 they have organised their own study and developed new indicators as shown in 2.3.2. According to their web page, OECD is “one of the world’s largest and most reliable sources of comparable statistics and economic and social data”, and also “one of the world’s largest publishers in the fields of economics and public policy”<sup>83</sup>. The organisation has also become central when it comes to production and publishing of educational statistics through the PISA reports as well as through *Education at a Glance* and *Education Policy Analysis*. The two latter publications are built on national data from the OECD member countries gathered in the UNESCO/OECD/EUROSTAT<sup>84</sup> (UOE) database<sup>85</sup>. These publications are seen as “a prime vehicle for disseminating the Organisation's intellectual output, both on paper and online”<sup>86</sup>,

---

<sup>82</sup> [http://www.oecd.org/pages/0,3417,en\\_36734052\\_36734103\\_1\\_1\\_1\\_1\\_1.00.html](http://www.oecd.org/pages/0,3417,en_36734052_36734103_1_1_1_1_1.00.html) (accessed 04.03.09).

<sup>83</sup> [http://www.oecd.org/pages/0,3417,en\\_36734052\\_36734103\\_1\\_1\\_1\\_1\\_1.00.html](http://www.oecd.org/pages/0,3417,en_36734052_36734103_1_1_1_1_1.00.html) (accessed 03.03.09).

<sup>84</sup> EU has its own statistical office, EUROSTATE, “to collect community-wide data on education inputs, processes and outcomes, and to present this information in a comparative framework based on a set of education indicators (Husén & Tuijnman 1994: 13).

<sup>85</sup> [www.oecd.org/education/database](http://www.oecd.org/education/database) (accessed 30.11.08).

<sup>86</sup> [http://www.oecd.org/pages/0,3417,en\\_36734052\\_36734103\\_1\\_1\\_1\\_1\\_1.00.html](http://www.oecd.org/pages/0,3417,en_36734052_36734103_1_1_1_1_1.00.html) (accessed 03.03.09).

and have been of importance in the member countries when educational policy has been shaped. Since more and more countries sign for PISA, the study gets influential also outside the member countries of OECD.

The organisation also runs *Centre for Educational Research and Innovation* (CERI). When reading the presentation on the web site, it is difficult not to compare it with the presentation of a think tank: “CERI has a particular concern with emerging trends and issues, futures thinking in schools and universities. We often have a longer time frame than most work, typically aiming to set an agenda for the next 5-10 years or longer”<sup>87</sup>.

It is important to remember the dual role of OECD; the role as an actor in the field of economy as well as an actor in the field of education.

Both individuals and countries benefit from education. For individuals, the potential benefits lie in general quality of life and in the economic returns of sustained, satisfying employment. For countries, the potential benefits lie in economic growth and the development of shared values that underpin social cohesion<sup>88</sup>.

Moreover it is important to see how the different roles intermingle: “No aspect of education systems – from financing to forms of selection, from pedagogy to questions of management – is spared the critical scrutiny of governments committed to market-driven change” (Jones *et.al.* 2008: 5). That is due to “the interaction of the programmes of national governments with the work of international organisations – the European Union and the OECD in particular” (ibid.: 5). In the case of the EU it is worth mentioning that their interest in education is quite new. The tendency to look at education as an economic tool has become particularly clear during the development of the Lisbon Strategy, where the goal to make EU “the most competitive economy in the world” was articulated<sup>89</sup> (Jones *et.al.* 2008). Hatcher (2009: 1) puts it this way:

In 2000, the European Council meeting in Lisbon decided that the principal objective of the policy of the Union with respect to education was the production of human capital for a competitive economy, and that policy is reshaping the education system of every country in the EU.

---

<sup>87</sup> [http://www.oecd.org/about/0,3347,en\\_2649\\_35845581\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/about/0,3347,en_2649_35845581_1_1_1_1_1,00.html) (assessed 23.04.09).

<sup>88</sup> [http://www.oecd.org/topic/0,3373,en\\_2649\\_37455\\_1\\_1\\_1\\_1\\_37455,00.html](http://www.oecd.org/topic/0,3373,en_2649_37455_1_1_1_1_37455,00.html) (accessed 30.11.08).

<sup>89</sup> [http://europa.eu/scadplus/glossary/lisbon\\_strategy\\_en.htm](http://europa.eu/scadplus/glossary/lisbon_strategy_en.htm) (accessed 01.12.08).



### 5.3.2 What kind of qualifications should be “produced”?

What kind of qualifications does the working life of today need? What kind of qualifications are emphasised in the OECD documents, and what kinds of qualifications are produced? In this section I will use OECD documents, the qualification framework and the human capital theory in the discussion of these questions.

#### *Human capital in the knowledge society*

*Human capital theory* reflects the “direct and functional relationship between education and development” (Kubow & Fossum 2007: 41). Education should improve the workforce according to the needs of the industry and contribute to the increase of the productive capacity and economic growth<sup>90</sup> (ibid.). EU, as quoted earlier and also OECD view the production of human capital to be a main objective.

Human capital plays an important role in the process of economic growth and individuals' labour market outcomes are linked to their educational attainment. The OECD reviews policies that influence the incentives to invest in human capital and the efficiency of the provision of education services<sup>91</sup>.

The production of human capital is needed in the *knowledge society*; where the length and the complexity of education has become a central issue (Jones *et.al.* 2008). Another concept is the *knowledge economy*, “where ‘knowledge’ takes over from ‘production’ as the key driver and basis of economy prosperity” (Dale 2005: 146). The human capital theory and the idea of a knowledge society and economy might, however, fit the ideas and the development stage of some countries and areas better than others. The human capital theory “reflects structural-functionalism’s fundamental affinity with economic liberalism embraced in the progressive democracies of the industrialized West” (Fägerlind & Saha 1989 in Kubow & Fossum 2007: 41).

Through PISA the OECD intends to prepare the countries for the knowledge society by providing them with comparable international analyses that “can extend and enrich the national picture by establishing the levels of performance being achieved by students in other

---

<sup>90</sup> See Table 3.1 and 3.2

<sup>91</sup> [http://www.oecd.org/departement/0,3355,en\\_2649\\_34605\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/departement/0,3355,en_2649_34605_1_1_1_1_1,00.html) (assessed 25.04.09).

countries and by providing a larger context within which to interpret national results” (OECD 1999:7). Through PISA the western ideas about human capital in a knowledge society are spread, but how do the theories fit in practice?

### *Today's reality*

In a developed country like the US the demand for non-routine analytic and interactive workers has increased since 1960, while the demand for routine cognitive and none routine manual labour force has decreased since respectively 1970 and 1960. However the demand for routine manual workers shows only a slightly decrease (OECD 2007: 35). British research reveals that there “has been an increase in demand for skills at the higher end of the labour market, and that the number of unskilled jobs (...) far exceeds the number of low-skilled workers” (Jones *et.al.* 2008: 41). More and more workers, but still only four out of ten in Norway in 2008 were employed in occupations that demanded higher tertiary education<sup>92</sup>. Hence, in a developed country like Norway, we still have a majority of workplaces that do not require tertiary education. Even the OECD (2007b: 35) admits that the drastic decrease in the need for manual workers as often claimed is exaggerated: The organisation

recognises that it would be ‘tempting’ to believe that ‘everybody can now participate in the new economy’ but feels obliged to temper this optimism: ‘there are many jobs located outside the most dynamic and highly qualified sectors. It is reasonable to suppose that there will continue to exist marked differences in occupational demand of qualification...(and) expansion is often accompanied by an increase in the proportion of non-qualified to qualified staff’(OECD 2001 in Jones *et.al.* 2008: 41).

There will always be a need for manual workers also in so called developed countries; those driving taxis and busses, those who clean our buildings and serve us in shops, restaurants, hairdressers and other places. At the present we have a stratified workforce in the *north* as well as in the *south*, however there is also stratification between the *north* and the *south*. In countries like the US, the UK and Norway there are more non-routine analytic and interactive workers than in less developed countries. Some countries in the *south* are by now taking over the routine cognitive work (OECD 2007b), India is being one of them. Routine

---

<sup>92</sup> According to *Labour Force Survey – data on occupations 2008* done by Statistics Norway’s labour Force Survey of <http://www.ssb.no/yrkeaku/> (accessed 25.04.08).

---

and non-routine manual work still dominates among work opportunities in most countries in the *south*.

If the rhetoric of the market liberalism is followed, it is the market that decides what kind of qualifications to produce. If there is a market for a certain amount of some particular skills, then we have to produce just such an amount. To produce more, would be inefficient. Thus the following quotation from the OECD is interesting. They argue that only the minority of students will manage to “enter a career in the dynamic sector of the new economy” and “that education programmes cannot be designed as if the entire school population were going to reach their higher levels” (OECD 2001 in Jones *et.al.* 2008: 41-42).

In chapter III I showed that the development of a neoliberal economy has been a process towards a more stratified world; the gap between those who have and those who have not has been widening, both between and within countries. I have tried to show that the development of the economy and the education has been quite parallel. Why should not the development within the education sector lead to the same type of divide? What if the one size fits all rhetoric in fact stratifies? Hill (2006: 45) based on 11 case studies from countries all over the world, concludes that “[l]iberalization is making provision of education services more unequal and selective rather than universal.(...) [offering] good quality schooling for the rich and the middle classes and poor quality schooling – or none – for the poor”. What if a stratified workforce is what is needed for some to become effective in the competitive world economy? What if the present development in reality promotes such a development?

*What kinds of qualifications are produced?*

According to the *qualification framework* (Masuch 1973 in Brock-Utne 2007a: 490), the educational system trains three different types of skills: *proficiency, adaptility and creative qualifications* (cf. 3.1.5.). What kind of proficiency level the labour force is trained for may vary according to the general development level of a society. Whereas an agricultural society might need a labour force with special proficiency qualifications connected to certain vocational skills, a highly industrialised society will have an increasing need for a labour force able to learn over and over again, and thus need general proficiency qualifications (Brock-Utne 2007a).

According to the rhetoric of OECD/PISA as shown in chapter II, the needs of general proficiency, as well as creative qualifications are outspoken. Seen at the backdrop of the reality of today one might ask if this is what really is promoted.

Workers in knowledge based societies need to be capable to update their knowledge, but not all of them need a high level of education, as seen earlier in this section. According to Jones *et.al.* (2008), “[a]daptability is certainly in demand; education, much less so: there is a distinction between the importance of innovatory capacity – as an arm of competition – and the general level of education called for by the knowledge society (ibid.: 41)”. Adaptability qualifications like diligence, perseverance, obedience, sense of duty and indifference might all be keys in a stratified society and labour market, whereas critical sense and independent thoughts like the creative qualifications nourish, might be dangerous if it is mastered by the many. Brock-Utne (2007a) has questioned what kind of qualifications the school system of Tanzania produces. This is a question that should be looked into in more countries.

### 5.3.3 Business models as seen in education

Business can make profit in and on schools, as it can do within other sectors. “Business is driven to seek new potential markets, and the public sector, including education, is seen as one of the few major areas still left to colonise” (Hatcher 2002:1). In short, business may involve itself in schooling on all levels; on international-, national-, municipal- and school level. Business may enter school systems and schools as sponsors, consultants and owners. They may provide books, technical equipment, leader- and teacher training and even teaching. However, since education traditionally has been a non profit public area, a change toward another way of thinking has to be prepared for.

For an educational market to work, it is essential to create competition within education and to change the conception of education as a right toward a market-oriented conception of education as a commodity. Comparable tests, results that are made public and free choice of schools are keywords. Comparable tests may enforce a more centralised curriculum, nationally, regionally and internationally and thus weaken the national sovereignty and lead to a more instrumental view of knowledge, to a more uniform education worldwide and to more privatisation. When standardised tests are implemented, results are published, students

---

have free choice of which school to attend and the money follows the students both competition and markets are created.

Different countries have taken on business models in education to different degrees. England<sup>93</sup> (Jones *et.al.* 2008) as well as the US (Apple 2006, Kohn & Shannon 2002) are countries where such models are most visible. In Norway they are more distinct within the health care system than within the education sector, although they are visible here as well. In this section I will use the structure of Hatcher (2009)<sup>94</sup> when I briefly discuss five ways the world of business might enter the school system. Hatcher uses examples from England; I will use examples from my own experiences in Oslo as a shop steward in the early 1990s and in mid 2000s.

*A market where parents choose schools and schools compete for students*

This first point is about creating a market where markets never where, quasi-markets. The idea of creating them stems from the private economy, where markets always have been a key. Educational markets answer to the demand of free choice of schools, which again demand decentralisation of power, a system of capitation grant or vouchers where the money follows the students, and the results on standardised tests are made public.

When I was a shop steward in the early 1990s, we had a centralised system in Oslo where the frames were negotiated at central level. There were fixed school borders, students were assigned to their neighbouring school, and there was an upper limit for students in a class. 28 in the case of primary, and 30 in the lower secondary compulsory schools. Allotment was given according to the number of classes and the real need for extra support to some students. We did not have any standardised tests. Emphasis was on school input. There were only a handful of religious and alternative pedagogy private schools.

In the mid 2000s there was a change towards more decentralisation. At least there was a decentralisation of responsibility if not of real power. The school borders were gone, there were no fixed limits for students in a class and allotment was given according to the number

---

<sup>93</sup> England, and not UK as a whole since Wales, Scotland and Northern Ireland have had a slightly different development in recent years (Hatcher 2009).

<sup>94</sup> The headlines in this section is quotations from (Hatcher 2009: 1,2,3,4,5), which means that the first quotation is taken from page 1, the next from page 2 etc.

of students enrolled. The results on the central tests, the Oslo tests and the National tests, were published. Emphasis was now on schools output. There were still very few private schools, since the election of the *centre/left* government in 2005 put an end to the former *centre/right* policy to allow private - so called free-schools - in Norway.

### *Regulation of teaching and learning*

A contradiction that has become visible is a simultaneous demand for decentralisation and centralisation. Decentralisation “may be presented as a democratising reform, but it is also one which is in most cases connected to marketisation” (Greens in Jones *et.al.* 2008: 57). According to Apple (2006: 63) the decentralisation tendency stems from the neoliberal vision of quasi markets, while the centralisation is caused by the neoconservative “pressure to regulate content and behaviour through such things as national curricula, national standards, and national systems of assessments”. The demand from the neoconservatives comes from the world of business, and is according to Hatcher (2009: 2) the result of “a culture of performativity, based on targets, contracts, evaluation, and a system of rewards and sanctions, designed to micro-manage the working lives and professional identities of teachers”.

The most striking difference from the early 1990s to the mid 2000s was the amount of formulas and questionnaires of all sorts we had to fill in. Earlier we discussed and formulated year plans, project plans, and week plans according to the curriculum, the visions and goals of the school as well as the need of our students. In the mid 2000, the goals and visions were formulated by the school authorities of Oslo; we only had to fill in forms. The last time I was a part of this process, we filled in the form with words. The school year of 2008/2009 I was told by a headmaster that they are now filled in solely with numbers.

### *A business- friendly curriculum*

Hatcher (2009) argues that emphasis on basic skills, information technology and business values as well as curriculum differentiation provides a business friendly curriculum.

Basic skills are easy to test, and can be seen as a premise for the growth of standardised tests, those again are a huge income source for business as will be discussed later in this section. Enormous private profits are made from the introduction of ICT in schools all over the world, both from the cost of the computers and software itself and from the costs of the

maintenance and updating of the equipment. “The development of information technology, especially the internet, is the largest and fastest growing sector of the new education-for-profit industry” (Hatcher 2002:12). Curriculum differentiations, that have been developed in England by offering an “impoverished curriculum for ‘non-academic’ students” according to Hatcher (2009: 3-4) “reinforces patterns of social inequality, and helps to create a stratified future workforce corresponding roughly to the differentiation requirements of the economy”. Although never mentioned in the OECD papers, this is what might become the reality when education no longer is seen as a right for the students and a good in itself, but only a commodity designed according to the need of industry.

#### *The school management of schools and teachers*

“The creation of a culture of performativity requires the transformation of the workforce and the creation of a range of new professional identities, from head-teachers to the classroom teachers” (Hatcher 2009: 4). As a shop steward I experienced the transformation of the role of the head-teacher, from the first among equals, towards the extended arm of the school authorities. I also witnessed how the control mechanisms were put in place to make head-teachers and teachers accountable. From 2003 the teachers’ salaries in Norway have been negotiated at municipality level (EI 2004 in Hill 2006). In Oslo this led to salary negotiations at school level and unpredictable salary differences, since the criterion for the rise in salary has never become quite clear. Moreover it led to a more competitive workplace for the teachers, who in some cases have felt that the salary was linked to pupils’ outcome. In England the transformation of workforce has led to the break of “the monopoly of qualified teachers by allowing other categories of school workers to carry out their jobs” (Hatcher 2009: 4).

#### *The role of the private sector*

In England the role of the private sector has expanded since the Labour government took over in 1997 (Hatcher 2009: 5). What Margaret Thatcher did not dare to do, Blair did (Jones *et.al* 2008). “Almost every major government policy initiative in the education system has relied on private companies to translate it into practise” (Hatcher 2009: 5). In England private companies provide teacher training and school inspectors, they run schools on non-profit and for-profit basis, and they are heavily involved in the production of all kind of resources needed in an output based education policy. In the latter case this applies for

student assessments as well as the “training for teachers in the national literacy and numeracy strategies to setting up the system of performance-related pay for teachers”(Hatcher 2009: 5).

In the US private companies do not only produce all the tests and test materials, they also score them. Kohn (2002) describes how five companies develop and/or score nearly all tests used in the US. The correcting is mainly done by non academic personnel on temporary contracts.

Business also enters into the schools in the US through advertising, where the goal seems to make students become consumers as early as possible (Apple 2006). The worst example is the story about Channel One. Each classroom is offered a satellite dish, two VCRs, television monitors and a news program. To receive this, the schools have to guarantee that students watch Channel One and its mandatory advertisements every day. As much as 40 % of middle and secondary schools in the US have signed a contract with Channel One (Apple 2006: 35). Baker (2002) views this as a major attack from business on vulnerable young people who are forced to view advertisements in classroom every day. He has calculated that Channel One takes up six or seven schooldays pr. year.

In Norway we only have a few non-profit private schools and there is an awareness not to bring advertisement into schools. We have always had private publishing houses. However, a growing emphasis on school outcomes opens up for private sector here as everywhere.

## 5.4 What kind of wheel, what kind of hub?

This chapter has been dealing with the neoliberalisation process within education, the process that also is called the *marketisation of education*. I use the metaphor *The Wheel* on the same process, thus I have experienced the present development within the education sector as evolving, from a tiny movement in the beginning until the present where “everybody”, the *right* and the *left*, move in the same direction at high speed.

I have seen the process from a historical, political, ideological and to a certain extent theoretical and practical perspective. In the next chapter I will use PISA to show what kind of knowledge the PISA assessments require from the students, what kind of knowledge they



---

produces and what kind of role standardised tests more general might have in the marketisation process. I will look at the process from a more theoretical; epistemological, ontological and methodological perspective, without losing sight of the other perspectives. I call the chapter *The Hub*, a metaphor I use on the standardised tests when I ask if they might be possible actors in the marketisation process; if they might be seen as *The Hub in the Wheel*.

## 6. The “hub”

*Measuring Student Knowledge* (OECD 1999: 9) presented the first PISA study. Then the participating countries represented “more than a quarter of the world population, more than that covered in any international educational assessment to date”. In PISA 2009 the number of countries has more than doubled, and its influence most probably as well.

“PISA is a gold mine of information” Harry A. Patrinos from the World Bank argued at CIES 2009<sup>95</sup>, and when reading the PISA documents it is easy to get impressed. Both their data presentations and analyses are easy accessible online. Although PISA produces complicated comparisons between multiple factors, what mainly is used by the media and the politicians, are the most simple league tables that compare students’ proficiency level in the different OECD countries. Those league tables are based on the cognitive part of PISA. To quote Langfeldt (2007: 225), they have been “the spearhead of PISA in attracting public interest”, and as I see it, the league tables have become actors in their own right. It is PISA as a political actor that has attracted my interest. Therefore it is those *most simple league tables* that compare students’ cognitive proficiency level across nations, that is my point of departure in the discussion in this chapter. Not the multitude of information delivered by PISA.

This chapter has three main sectors. First I will discuss the issue of setting standards, then I will question whose knowledge that counts. Thereafter I will look into what kind of knowledge that counts.

### 6.1 Setting standards

Standardised tests require that someone define standards, construct tests according to them and use them while scoring. According to Cummings (2003: 165), some have no problems to set standards in education, others argue that cognitive outcome should not be attached to

---

<sup>95</sup> At the meeting *Global Monitoring, International Aid, and Country Development* 24.03.09

standards and that such results can not justify any comparison of schools. In between we find those taking an intermediate position.

### 6.1.1 Technical challenges

#### *Difficulties*

Regardless of which position to take, scholars realise the difficulties of setting educational standards. Husén & Tuijnman (1994: 3) argue that it is much more difficult than in the case of physical standards, such as length or weight. They are invariant, while the “main problem of standard setting in education is that the underlying goals are subjectively derived and subject to change with the passing of time”. They further argue that “there is no firm agreement on what such standards are, how they are best set, and what their relationship is or should be to the content items used in scoring students performance”(ibid.:2-3). PISA *insiders* are also concerned. Procedures to guide the construction, the scoring and the coding of test items have been developed in details (OECD 2009).

#### *Cut-off score*

Gonzales & Beaton (1994: 171) question the cut-off score on the measurement scale. “[T]he standard to reach is translated into a cut score on the measure. On one side of the cut-score those that have reached the standards are placed”. They argue that the “procedure used in the setting of standards and the selection of appropriate cut scores have been often criticized” and that the setting of cut scores are “based on or determined by individual preference or convenience rather than by necessity or the intrinsic nature of something” (ibid.:176):

[I]t should be noted that the decision to set a cut score is principally a political matter. The decision is always an arbitrary one because it is ultimately based on value judgment – albeit judgments that are informed by technical considerations. The decision to set a cut score for a standard is not fixed within law but depends on the individuals who select the methods, develop the items, design the scale, and who eventually determine the cut score itself (ibid.: 188-189).

#### *What kind of standards?*

Important to have in mind while analysing PISA is awareness of what kind of standards the students’ achievement are measured against. The four different main assessment types, those *testing knowledge, skills, values/attitudes/convictions* and *aptitudes* have to be seen against

*opportunity to learn standards, content standards, and performance standards*<sup>96</sup>, as well as their interconnections. “[C]ontent standards and performance standards are interdependent, and ( ) they are conditioned by *opportunity to learn* criteria” (Husén & Tuijnman 1994: 3).

#### *Not only a technical exercise*

Phillips (1994: 209-210) delivers the following expressive description of the process of setting standards like this:

Setting performance standards is a little like cooking soup. The recipe calls for a delicious blend of art and science. Experience and experimentation will improve the flavour, which is influenced by tradition and cultural preferences. However, no matter how well the recipe is followed, and no matter how pleased the cook is with the outcome, there will be those who do not like the taste. Public policies often require that standards be established out of practical necessity even though the activity does not have a firm scientific basis. Standard setting is best accomplished in an environment in which policymakers with a vision of the public good work collaboratively with technicians willing to take a risk.

He indicates that there are more into the procedure of setting standards than technical challenges. Husén & Tuijnman (1994: 16-17) make this issue more clear:

Whereas the goals for education are mostly determined at a political level, their translation into concrete standards and procedures involves as much a political process. Standards are informed by public and professional discourse, by the judgment of influential actors, and by international benchmarks derived from international studies. Accordingly, policy considerations, scientific and technical considerations, and practical issues jointly influence the determination of standards

### 6.1.2 A question about will and power

#### *The will to use standards*

Regardless of the difficulties in setting standards, we have seen an emerging will to use standards and standardised tests in the production of knowledge about education systems all over the world. In chapter II we saw how this will developed from almost rejection in the 1970s, at least in most of the European countries, to almost an embracing at the present. As a consequence we have seen a growing test culture that answer to strong demands for quantification in the present social sciences (Kvale1996), where preparation, implementation and correcting of standardised, quantifiable tests might occupy as much as 10 % of instructional time and 10% of educational expenditures in some systems (Cummings 2003:

---

<sup>96</sup> The different assessment types are explained and the different standards are mentioned in chapter I.

193). While looking at the rise in the amount of tests as showed in chapter II and in appendixes, my hunch is that more and more countries might be on their way to reach the same level of expenditure by 2009.

Important to remember, however, is that knowledge derived from standardised tests is conditioned by epistemological, ontological, methodological – and human - limitations. Why is there a demand for such knowledge and a willingness to accept the growing use of it?

Kvale (1996: 67) makes the following comment to this question:

There may be an *ontological* assumption that the social world is basically a mathematically ordered universe in which everything that exist, exists in number form; and, accordingly, the objective data of science of the social world must be seen quantitatively. There may also be an *epistemological* demand that research data should be quantitative in order to be commensurable across theories. There may further be a *technical* interest in quantification, in that statistical techniques are powerful tools for handling large amounts of data. The demand for quantification may also stem from the anticipated audience of a research report, such as a dissertation committee, the scientific or public community, or a government agency. The use of numbers may be *rhetorical* here; when it comes to convincing a modern audience, the hard quantified facts may appear more trustworthy than qualitative description and interpretations.

Jones (2007a: 331) deepens the answer of the above question:

[T]he issue at hand is more a matter of how interests and motivations in educational theory, policy and practice intersect with and are driven by powerful global constructions of educational values and techniques. Thus, despite the persistence of local uniqueness, education around the world is seen to become increasingly standardized. Identical reform agenda emerge, in the most disparate places, rich and poor education systems being moulded and reshaped in accordance with uniform measures, Global hierarchies of knowledge emerge, giving preferred status to some lines of thought and methods over others. Global systems of evaluation are adopted, testing the extent to which national systems are conforming to privileged global standards and reform agenda.

*Whose standards?*

Those defining standards have a power. Who is in the position to define standards and thus hold this power? Are the possibilities to define standards equally divided between the *lowers* and the *uppers*, or between the *south* and the *north*?

PISA is an OECD project. Their secretariat has the overall managerial responsibility for the program and “produce the indicators, and analyse and prepare the international reports and publications in co-operation with the PISA consortium and in close consultation with the Member countries at the policy level” (OECD 1999: 17). Although they emphasise the

collaborative aspect, one has to remember their ultimate goal and see the PISA project in this light.

The overall perspective of the OECD is concerned with the market economy and growth in free trade world. All policy advice they provide is certainly coloured by such underlying value commitments. Hence, the agenda of the OECD (and PISA) does not necessarily coincide with the concerns of many educators (or other citizens, for that matter). The concerns of PISA are not about “Bildung” or liberal education, not about solidarity with the poor, not about sustainable development etc. - but about skills and competencies that can promote the economic goals of the OECD (Sjøberg 2007: 209).

OECD has always been a major market liberal actor. In chapter III I showed the global impact of market liberal economy standards. Now it is timely to ask about the global impact of the educational standards offered by PISA. The standards of PISA are very powerful tools. They are used as criteria of success or failure of a country’s educational system. Do they pay respect to the different cultures and countries in which they are used? Do they test what they say they test - is PISA according to PISA? Or is it possible to find a hidden agenda behind all the nicely formulated documents and graphs?

## 6.2 Whose knowledge counts?

Whose knowledge counts is connected to whose standards count. In the document *Measuring Student Knowledge and Skills* (OECD1999: 17) it is stated that all the member states are involved in decision processes to assure that

- the instruments<sup>97</sup> are internationally valid and take into account the cultural and educational context in OECD Member countries;
- the assessment materials have strong measurement properties; and that
- the instruments place an emphasis on authenticity and educational validity.

Nowadays all the participating countries are suggested to have a say. “PISA is a collaborative effort by the participating countries, and guided by their governments on the basis of shared policy-driven interests. Representatives of each country form the PISA Governing Board which decide on the assessment and reporting of results in PISA” (OECD 2007b: 3).

---

<sup>97</sup> I interpret the term *instruments* to be the test items.

Nonetheless, my issue in this section is to question if, in spite of the efforts, there are biases connected to the test items of PISA, biases that give preferences to some and thus value some kind of knowledge over other.

### 6.2.1 The test items of PISA

Who makes the test items? For which purpose? In what way? The process of setting standards is determined by technical as well as more value laden aspects. The technical part is meant to counter all thinkable challenges. Nevertheless critics of the technical guidelines exist, as I will show. The main critique of the test items is, however, connected to more fundamental questions. This critique challenges more value laden and fundamental issues. Since this is a critical discourse analysis, I will emphasise on the more fundamental critique.

#### *Cultural biases?*

Our socio-cultural praxis sets a standard for our thoughts, Østerud (2006: 211-212) argues. He asks if PISA test items may be more in line with some cultures than others. Langfeldt (2007: 232) states that “pupils from different countries will have systematically different chances to perform equally well”. He becomes more specific when he asks if “the real world ambition of PISA refers to a world shared by all” (...) or if “the concept ‘real world’ competence (...) [is] an Anglo-phone preference”. Wuttke (2007: 258) argues that “[C]luster analyses invariably show that student behaviour is most similar in countries that share both language and cultural heritage”, an interesting argument when the following information is added:

A majority of test items comes from English-speaking countries; the other items were translated into English before they were streamlined by ‘professional item writers’. If there is cultural bias, it is clearly in favour of the English-speaking countries. This makes it difficult to separate it from the translation bias, which acts in the same directions (ibid.:257-258).

In spite of the collaborative efforts of PISA, it might be important to remember that it is the consortium in collaboration with expert panels that set the standards and inform about how to code and interpret the students’ assessments (OECD 2009). The consortium has been lead by ACER from the very beginning (OECD 1999). Up to 2006, members of the consortium represented only western culture (OECD 2009). If the reason was to better reflect the different cultures of the countries participating in PISA, one might ask if the restructuring of

the consortium in 2006 managed to make much of a difference. Except for the present Japanese participation, it is questionable whereas institutes in Australia, the Netherlands, Norway and Germany really represent culture diversity although their educational systems differ.

When looking at education as highly contextual, it might be difficult to believe that cultural neutral items are possible to construct:

As long as learning, as well as all mental and physical actions is contextual, learning is only possible to understand and describe as part of continuous, social processes. It will not be possible to formulate texts that can yield justice to all the different countries [participating in PISA] (Østerud 2006: 213).

In the following I will exemplify some of the problems that can occur when standards is supposed to apply to a multitude of students cross countries, languages and school systems.

#### *Authentic texts?*

One of the demands is that the test items should be authentic. This is met by using texts that have been on print in one of the participating countries. When it comes to authentic texts four main questions should be asked:

Is it realistic to find authentic texts suitable in all countries participating in PISA? Does *anyone* really believe that it is possible to find texts that are experienced as authentic for 15 year old students in countries like Norway, Colombia, US, Kyrgyz Republic, Qatar, Thailand, Tunisia and New Zealand<sup>98</sup>? Will some students experience them as more authentic than other students? Are test items used in 2000 still “authentic” in 2009? Do texts remain authentic after being translated? Sjøberg (2007: 220) argues that the Norwegian translation of test item S128 *Cloning* from 2003<sup>99</sup> is a good example of an authentic text that has lost its authenticity as well as its scientific content during the process of translation.

---

<sup>98</sup> Countries chosen from different continents and cultures taken from the list showing participating countries in 2009: [http://www.pisa.oecd.org/document/4/0,3343,en\\_32252351\\_32236225\\_39758660\\_1\\_1\\_1\\_1,00.html](http://www.pisa.oecd.org/document/4/0,3343,en_32252351_32236225_39758660_1_1_1_1,00.html) (accessed 12.03.09). It is important to add that only OECD countries are presented in the league tables, so all of those countries mentioned here will not be compared. However here we talk about the problem with authentic texts, not comparison.

<sup>99</sup> The few test items that have been freed are found here: [http://www.pisa.oecd.org/document/38/0,3343,en\\_32252351\\_32236173\\_34993126\\_1\\_1\\_1\\_1,00.html](http://www.pisa.oecd.org/document/38/0,3343,en_32252351_32236173_34993126_1_1_1_1,00.html) (accessed 01.04.09).



---

Dolin (2007: 101) states that “the more realistic a test is – the more it is designed to reflect actual everyday situations – the less it makes sense to compile a globally comparable test”.

*Whose language counts?*

Translating procedures might act differently in different cultural and linguistic contexts. Puchhammer (2007) has compared the German and English versions of 24 released PISA mathematic items, and managed to show how translation might make a text more complicated. He found that the German texts are significantly longer, use more rare words<sup>100</sup> and contains more difficult grammar than the English. In addition more sentences contain subordinated clauses in German. Wutkke (2007: 257) argues “that the French texts contained on average 12% more words and 19% more letters” than the English. Hence the German and the French items take more time to read, something that obviously must alter the test scores<sup>101</sup>. These problems apply for all items, since all of them, regardless of what kind of literacy they tests, contain lots of words.

Language issues are not only a cross national issue. They might also occur within countries where parts of the population are minority speakers. One of the problems is connected to the translation problem mentioned above. Usually the most frequent words in a language is taught first, thus people with different mother tongues might get difficulties with PISA items in countries that use translated texts (Puchhammer 2007). Another problem is the one connected to the making of test items. According to Dolin (2007: 109) “it is impossible to avoid a certain amount of cultural bias. Test items that require the students to read between the lines in reference to cultural background knowledge are managed more easily by Danes than by Danish students from ethnic minorities”. Those two issues might be the reasons why the mean score of minority speaking students in Norway are lower that of majority speakers (Hvistendahl & Roe 2003).

---

<sup>100</sup> To find out if words are rare or common in a language, it is possible to use Zipf’s law that holds for most languages. While using this law, it is possible to link statistical frequency of word and rank frequency of occurrence. Built on this law, word frequency tables have been constructed in some languages, English, German and French are three of them. These tables rank the 10000 most used words in the respective languages (Puchhammer 2007: 133). These tables are used while comparing the different versions of PISA test items. “Although only a few words have been selected, the result is impressive” (ibid.: 135).

<sup>101</sup> Similar comparisons ought to be done between English and other languages to find out whether this is a common problem.

*Whose real life counts?*

PISA is not supposed to test school curricula, but assesses “knowledge and skills that are essential for full participation in society”<sup>102</sup>. Is real life the same for students in all the countries participating in PISA? Sjøberg (2007: 214) notes: “Real life is different in different countries”. For students the life they live is connected to the school system as well as the society they are raised in. Some school systems raise students to be competitive, conformist and disciplined, others emphasise on cooperation, creativity and initiative. The way students are raised might become decisive in their meeting with PISA.

Sjøberg (2007) asks whether Norwegian youth has the willingness and motivation to do their best in such a test that does not mean anything to them personally. They will not be graded, and the test contains items that are long, boring and not always well formulated. At least this holds true for the Norwegian translations, according to Sjøberg (2007). Moreover he compares the Norwegian students with those in Taiwan and Singapore who enter the test premises in the spirit of competing in a world championship. 75% of those participated in the study conducted by Union of Education Norway mentioned in chapter V supports his doubt (Sjøberg 2009: 26).

I have only seen the Norwegian translation in Sjøberg (2006: 199) and the one used in the Oslo test in 2003. I have however looked into the English version of the few test items that have been freed from 2000, 2003 and 2006. My reaction is that they are more school related in the traditional meaning of the concept than real life connected, and hence I have difficulties to understand how they can be perceived as expression of something of current interest by 15 year old Norwegian students. Simultaneously they represent a form that is unfamiliar for Norwegian students; a mixture of reading, science and mathematic literacy test items in each booklet. Only 22% of the participants in the Union of Education Norway’s study mentioned earlier answered that the Norwegian students are familiar with this type of test (Sjøberg 2009: 26). The mixture most probably puzzles them and makes them anxious, as also the time limitation is likely to do.

---

<sup>102</sup> [http://www.oecd.org/pages/0,3417,en\\_32252351\\_32235918\\_1\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/pages/0,3417,en_32252351_32235918_1_1_1_1_1_1,00.html) (accessed 06.01.09).

---

*Whose form and habits counts?*

“Coping with life in modern societies requires a range of competencies and skills that cannot possibly be measured by test items of the PISA units’ format” (Sjøberg 2007). Sjøberg is particularly concerned about science, where the understanding of issues is closely connected to handling objects. Dolin (2007: 107-108) asks “whether one can reasonably claim that sitting with a paper and pencil and (casually) answering questions about imaginary situations has anything to do with competencies in the sense that we normally understand them”.

Research has shown that students experienced in multiple-choice questions solve them quicker than those not familiar with this form. There are also country specific reaction forms when it comes to timing, tactics and fatigue (Wuttke 2007).

*Whose experience counts?*

In 2006, the school I worked at took part in PISA and five of my students were sampled to take the test together with students from other classes. During the test I entered into the room where the students wrote their answers to check if one of my students struggling with heavy dyslexia had turned up. He was really upset about taking part in the test, since he knew he would not succeed in a pen and paper test, while in a natural setting in the laboratory he always did quite well. He also managed well when he could use a computer, which he in fact had received for free because of his severe dyslectic handicap. To my big surprise entering the classroom was like entering into a hotel lobby. Many of the students did other things than answering the test and some were even talking to each other. I asked the teacher watching them what was going on, and he said that many had finished, that they had different booklets and hence talking did not matter.

After the test I talked to my students. Both the bright ones and those who were not that bright, told me that they got confused since they thought the test was supposed to ask questions about science, but that it also asked lots of other questions. The bright students were stressed about the time pressure as well as the noise from the students talking. Those not that bright told me that they did not understand a thing and that was the reason why they started to talk to each other, although they were not finished at all. The school in question is a multicultural school, and one of my students coming from the Middle East area was upset with the questionnaire asking all sorts of personal questions which she found inappropriate to

answer. All of them concluded that they actually did not care about the result after all, since the tests were not graded and they would not get them corrected.

The experience from my own students is not possible to generalise. They were only four, since my dyslectic student shirked school, according to my knowledge, for the first and last time in three years that day. However their voices are interesting to listen to. They give a mirror into how the PISA assessment is experienced, something important to investigate more systematically.

### *Whose ontology counts?*

Whose knowledge is the most important? The knowledge derived from lived life, or the knowledge that exists independent of social actors? The former position is *inductive*, or “bottom up” and is called *constructivism*. A pedagogy that builds on this ontological position will emphasise on students’ background, interests, cooperation, and processes. A research methodology building on the same position will value a qualitative approach showing experiences from lived life. Freire (1996) regarded the rote learning in *banking education* as oppressive and a way to reproduce inequality. The same thoughts are found in *Dependency theory*<sup>103</sup>. Freire (1985, 1996) also emphasised on the potential power of education to contribute to the end of oppression *if* the education is designed in a way that allows students to build on their preconditions, intrinsic motivation and develop their self-esteem. These are the same thoughts that are found in the *Liberation theory*.

In a period after WWII, those thoughts were visible in the discussion about what type of education to develop (Jones *et.al* 2008). They even became visible in the Norwegian curriculum plan of 1987 where the preconditions of students were emphasised as important to learning (Kirke- og Undervisningsdepartementet 1987, Øzerk 2006). Local syllabi and the rights of the minority speakers were emphasised. Øzerk (2006: 48) calls this emphasis an attempt to promote “inside-out” learning. Dolin (2007: 116) argues that “[i]n the Nordic countries, we have built up a view of knowledge in an educational context which attempts to combine the process-oriented view of knowledge expressed by constructivism with the more absolute view of knowledge expressed by science”.

---

<sup>103</sup> See table 3.1 for an overview of the Modernist theories, where the Dependency theory and the Liberation theory are two Marxist inspired theories.

---

Standardised tests do not fit into a constructivist way of thinking, because such tests set standards that define what type of knowledge that counts and hence are *deductive*, or “top-down”. The more students from multiple backgrounds who participate in a standardised test, the bigger the danger of alienation towards the test items and their embedded standards might become. The gap between the students precondition and the test item might become decisive. Moreover, whether students cope or not will depend on how their own experiences differ from *those* defining the standards. Those who set the standards are likely to define whose knowledge counts.

OECD/PISA (1999: 9) underlines that the assessments “will not primarily examine how well students have mastered the specific curriculum content. Rather their aim is at assessing the extent to which young people have acquired the wider knowledge and skills in these domains that they will need in adult life”. This definition removes the assessment from promoting rote learning and also banking education, thus “although specific knowledge acquisition is important in school learning, the application of that knowledge in adult life depends crucially on the individual’s acquisitions of broader concepts and skills” (ibid.). What is rhetorically promoted is the understanding and the ability to use knowledge. The problem is, as I see it, that OECD/PISA rhetoric does not match the reality. For the students to show how they are able to use the knowledge they have acquired, there are some preconditions which have to be met. If the test items are experienced as alien, if the translations of the test items make the test more difficult, if the school system emphasise on developing different values than those of PISA, the students might not be able to show the correct level of proficiency. Not only what type of knowledge that counts, but also *whose* values and knowledge that count, might become decisive for the results.

*Objectivism* is the ontological position counteracting the former. Within its deductive approach standardised tests fit very well, since knowledge exists independent of social actors; students just have to adapt to it. Not only knowledge might be standardised, but also values implying what the school is supposed to produce. The development of *human capital* to improve the workforce and the productive capacity, was one of the core stones in the structural-functionalist theory<sup>104</sup>, as well as it is in some theoretical school of thoughts within

---

<sup>104</sup> See table 3.1.

postmodernism and poststructuralism (Kubow & Fossum 2007). During the neoliberal offensive this theory has become the outspoken goal of education. In the explanation to why PISA does not test curriculum but mastery, the development of human capital in order to prepare students to become lifelong learners is explicit (OECD 1999). The concept of lifelong learning is one of the concepts that have been emptied with its original meaning and refilled to meet the neoliberal need of flexibility in order to contribute to a never ending economic growth (Jones *et.al.* 2008) The narrow emphasis on education as an economic investment in order to compete in the global market, has been criticised as a way to mainstream education globally, and PISA has been seen as an actor in such a development (Jones *et.al.* 2008, Sjøberg 2007).

A pedagogy that reflects the theories of Marxism that is showed in part of the work of Freire (Small 2005) and also seen in the Nordic countries is called *progressive pedagogy*. This pedagogy do not answer to the demand for economic growth, but aim to develop the students' personality and to promote critical thinking and thus might be seen as a threat to the hegemonic economic project of neoliberalism. Accordingly, a policy to counter a development in the direction of a more progressive pedagogy is understandable if the aim is to change the way people think, not the reality they live in.

## 6.3 What kind of knowledge counts?

This section has two levels. The first discusses what kind of knowledge and skills that counts for the students to master. The second discusses what kind of knowledge PISA produces; the knowledge that since the result of the first PISA study was published in 2001 has become the one that counts for politicians in so many countries.

### 6.3.1 What kind of knowledge and skills counts for the students?

The cognitive part of PISA is designed to assess knowledge and skills essential for participation in society (cf. 6.1.1.). Therefore PISA is supposed to assess the students according to performance standards, not according to content standard since it is supposedly not curriculum based, and not according to the opportunity to learn standards. Only the questionnaire might monitor the latter. However, Husén & Tuijnman (1994) underline the interconnections between the performance, content and opportunity to learn standards. Is it

possible to formulate authentic and cultural neutral test items that only assess knowledge and skills according to the performance standards? As shown in the section above, it seems difficult.

In addition to the more general problems attached to the test items as shown in 6.2.1, some question if PISA assesses knowledge according to content standards, others if PISA is an intelligence test.

TIMSS is curriculum based, PISA is not. Accordingly one expects the results on those two studies to vary. However, two Norwegian researchers, Lie & Olsen, presented a paper at a conference in 2007 showing a comparison of TIMSS and PISA science results from 22 countries. They found the correlation between test scores in the two studies to be 0,95 (in Dolin 2007: 113), causing questions whether the items of PISA are more connected to curriculum than intended. This correlation might also support the argument indicating that PISA is an intelligence test.

Wuttke (2007: 260) argues that the correlation between competences from different parts of the PISA booklet is high, and “between 75% (Greece) and 92% (Netherlands) of the total variance of students competences can be attributed to just one component” something that according to him makes it “hard not to make the connection to the *g* factor of cognitive psychology”<sup>105</sup>.

Questioning if PISA represents a real life challenge and offers authentic texts to the students, Dolin (2007: 113) utters “is it not the case that the more the test items and the test situation are shorn of their context and removed from ordinary everyday life, the more we tend to test levels of general intelligence”?

If cultural differences, translation, motivation, form and habits influence how the test items are experienced and mastered, the study might involve the testing of knowledge, skills, values, attitudes and aptitude according to performance, content as well as opportunity to learn standard. Then PISA has two main problems: Firstly, the study does not test what it is

---

<sup>105</sup> The *g* factor, also called the general factor. Spearman, the British psychologist active in the IEI inquiry “concluded that just two kinds of factors underlie all individual differences in test scores. The first and more important factor, which he labelled the ‘general factor’, or *g*, pervades performance on all tasks requiring intelligence. In other words, regardless of the task, if it requires intelligence, it requires *g*”. The quotation is taken from: <http://www.britannica.com/EBchecked/topic/228450/general-factor> (accessed 02.04.09).

supposed to test. Secondly there are uncertainties concerning what the study compare. The first problem is, according to Dolin (2007) as quoted below, related to the most fundamental question in a comparative test; the validity. The next is related to the reliability, both crucial demands for the generaliseability of the results in a comparative study.

### 6.3.2 What kind of knowledge is produced by PISA?

PISA is a quantitative comparative study. Such studies have to follow strict statistical rules. Central in quantitative research is to meet the demand for reliability and validity:

In a comparative test, reliability is crucial. Irrespective of what you measure, it must be done correctly. You must be certain that the various countries are appraised in the same way, so that their ranking in the final evaluation will not be open to question. Reliability-related problems include, for example sampling procedures and the scoring of responses. The most fundamental question, however, relate to the survey's validity - the extent to which the chosen design can measure what you are interested in. There is a gradual transition between problems of reliability and problems of validity, so the divisions between them are as much question of organisation as content (Dolin 2007: 98).

In this section I will look into the sampling and scoring procedures of PISA.

#### *Target population*

“The desired base PISA target population in each country consisted of 15-years-old students attending education institutions located within the country, in grade 7 and higher” (OECD 2009: 64). However, models of schooling are not universal<sup>106</sup>. Differences created by different systems in diverse countries, are not taken into consideration, such as

- variations of age groups at the end of compulsory schooling; in some countries students repeat or skip classes (Langfeldt 2007: 229)
- different drop out rates in different countries (Langfeldt 2007: 229, Wuttke 2007: 246)
- enrolment in preschool<sup>107</sup>
- the amount of instructional hours the target group has received during schooling (Langfeldt 2007:229)

---

<sup>106</sup> One of the main points underlined by Lawrence J. Saha from Australian National University when he criticised the PISA study in the meeting *Globalisation and comparative Global Pedagogies* 25.03.09 at CIES conference 2009.

<sup>107</sup> Joseph Zajda from Australian Catholic University at the meeting *Globalisation and comparative Global Pedagogies* 25.03.09 at CIES conference 2009, where he showed that the enrolment in preschool in Turkey is 5% and in France is 90%.



- some of the students are at the end of the compulsory education, other in the beginning of a new educational course equivalent to form 1<sup>108</sup>
- “[t]he 15-years-old international target population was slightly adapted to better fit the age structure of most of the Northern hemisphere countries” OECD (2009: 64).

All the points mention here, caused by structural diversity between different countries, make it problematic, if not to say unfair, to compare the PISA results of different countries. The heterogeneity in the target population is connected to sampling problems; hence it is a reliability problem.

### *Sampling procedures of PISA*

In comparative tests, sampling procedures that ensure the representability of the people involved is crucial to gain external validity (Vulliamy 1990). For a study to hold high external validity, the results must be generalizable. In the book *PISA according to PISA* (Hopmann, Brinek & Retzl 2007) several sampling problems are mentioned, and thus there is a question to what degree the study is external valid. In this section I will refer to the sampling problems that seem to be most important.

“To provide valid estimates of student achievement, the sample of students had to be selected using established and professionally recognised principles of scientific sampling, in a way that ensured representation of the full target population of 15-year-old-students” (OECD 2009: 65). Although the rules are clear and PISA offers detailed sampling guides, they are difficult to fulfil in some countries. Sampling is done in two stages<sup>109</sup>, within strata and within schools. The response rate varies in both stages. Prais (2007) mentions the difficulties of strata sampling in England and the following need for replacement schools. Wutkke (2007) discusses other countries, showing that response rate has been lower than required. He argues that low response rates both at strata and student level might cause considerable bias, since it is a tendency for “weaker” schools and students not to participate, something that might increase the measured national average.

---

<sup>108</sup> Henry M. Levin from Teacher College, Columbia University at the meeting *International Achievement Studies on National Education Policymaking* 22.03.09 at CIES conference 2009.

<sup>109</sup> In 2006 one country used a three stage sampling (OECD 2009: 64).

Even if in OECD (1999: 9) it is argued that the aim “is to assess the cumulative yield of education systems at an age where schooling is still largely universal”, not all countries participating in PISA, not even all the OECD countries, have full enrolment or enrol solely 15 years-old students in the school-year when PISA is conducted. Because students skip or repeat classes, in countries like France, Germany and Switzerland (Prais 2004 in Langfeldt 2007: 229) “[o]ften by the age of 15 hardly more than half of the pupils may be in class for pupils of that age”. If we look to Mexico, Turkey and many of the OECD partner countries “enrolment is less than 60%. Obviously, PISA says nothing about outcomes of the education systems in these countries” (Wuttke 2007: 246).

The relevant issue for a discussion of reliability is whether those who drop out have the same academic proficiency as those who stay in school. Arguing the case for better home background being decisive for academic achievement, a case being so well researched that it would be trite to mention evidence, one may well assume that the pupils quitting before the age of 15 as being unequal to the ones remaining in school. This leads to the conclusion that in a PISA context, some nations gain from the fact that as much as 60% of their classmates have left school before the age of 15 (Langfeldt 2007: 229).

Wuttke (2007) argues that different countries exclude more students of the target population than the 5% allowed and use different criteria for whom to exclude. An interesting piece of information is that Finland, the “gold winner” of the PISA league tables, excludes students with dyslexia. Denmark excludes those with dyscalculia and in Luxemburg immigrant students are excluded. Hörmann (2007) and also Wuttke (2007) discuss the exclusion of students with disabilities. In some countries students from vocational tracks are underrepresented, in other countries the technical difficult procedure of sampling leads to a target population above 100% (ibid.). Uneven gender sampling might also alter the country score. Except in Korea, where girls are underrepresented, the tendency is that girls are overrepresented. This can also alter the country score, since it is found some gender related score differences, both in connection to test items (OECD 2007a: 71&72).

Errors connected to strata and school sampling, drop outs, students that have skipped or repeated classes, exclusions of marginal groups or underrepresented groups make the study less reliable. Moreover it makes the countries less comparable. In fact high drop-out rate might end up as acting as a success criterion in PISA! The inclusion of fewer “weak” students in the sample might alter a country’s mean score and give it better placing on the league tables. Taking into account what impact those league tables have had, the sampling

errors should not be overlooked. They are the base on which the knowledge PISA produces comes from.

### *To score and code*

Quantitative research has to convert all data into numbers to be able to measure the outcome. The multiple choice items are “coded” by the students, but “45 % of the cognitive items required manual coding by trained coders” (OECD 2009: 111). How this is done is as important for the reliability as how the students are sampled:

It is crucial for comparability of results in a study such as PISA that students’ responses are scored uniformly from coder to coder and from country to country. Comprehensive criteria for coding, including many examples of acceptable and unacceptable responses, were prepared by the consortium and provided to NPMs [National Project Managers] in coding guides for each of science, reading and mathematics (ibid.).

At national level the National Project Managers had to select coders and train them. “It was not necessary for coders to have high level academic qualifications” (ibid), but they needed understanding of mid-secondary level science and mathematics and the test language, and also to how students at the age of 15 express themselves (See OECD 2009: 111 – 120 for more details).

In 2003, Oslo carried out the first Oslo test in reading competencies for standard 10 students, a test constructed in the same pattern as PISA, using PISA test items from 2000<sup>110</sup>. The idea was to gather the results from the different schools and make comparison between them. Teachers teaching the subject Norwegian in standard 10 should correct their pupils’ answers. To be able to code them correctly, all those teachers were gathered for one day to learn to code. I was one of them. Like in PISA, some answers had only two possible scores, false or correct, others had three possibilities; correct, partly correct or incorrect. We, the highly educated teachers in Norwegian language, were not at all that united when it came to what kind of answers to score as most correct or partly correct. When I later corrected the tests of my students, I met difficulties again. The manual and my judgment did not always fit,

---

<sup>110</sup> The report following this test is found here:

<http://www.utdanningsetaten.oslo.kommune.no/getfile.php/Utdanningsetaten/Internett/Dokumenter/rapport/satsningsomrxd/erx2002/lengre%20oppsummering%20av%20o-p-rapport.doc> (assessed 29.03.09).

The English version of test items used is found here: <http://www.pisa.oecd.org/dataoecd/44/62/33692744.pdf> (accessed 01.04.09). One of the items that caused lots of discussion was Reading Unit 8: 58-69; *The gift*.

moreover discussions between colleagues showed differences in how the students' answers as well as the manual were interpreted. Additionally the attitude of the teachers differed, some willing to interpret the pupils' answers as mildly as possible, others were strict. At the end of the day our class scores were compared with other class scores by our headmaster, and the scores of our school compared with those from other schools. Were they really comparable?

As in all standardised tests, the standard multiple choice items of PISA are automatically coded by students themselves, while the complex multiple choice items are coded by coders. They are easy to code since there is only one correct answer. Nevertheless to find the correct answer to a multiple choice task assumes a sort of common understanding between the one making the question and the one answering. Whether the banana is yellow or white depends on whether it is peeled or not and hence can cause different tick offs<sup>111</sup>. However the most severe problems are connected to the free response items. Was it possible for an unknown number of coders<sup>112</sup> from the 57 countries that participated in 2006 to code more than 400.000 assessments exactly the same way? If this was not possible, PISA 2006 had a reliability problem. If this in general is possible, there is a danger that the questions in the test items are cognitively too limited. In his book, *The Case Against Standardized Testing*, Kohn (2000) has a clear message; standardised tests do not ask the interesting and important questions since such questions are too difficult to code. Dolin (2007: 113) questions whether challenging test items are possible to use in comparative assessments:

The better the evaluation is at capturing complex skills, the more difficult it is to present the results in the form of simple, comparable data.

This brings us back to the traditional dilemma between undertaking an evaluation with a high degree of validity which is costly to carry out and which, because of its complexity, will have low reliability, and an evaluation of simple factors which is capable of measuring with high reliability, but in which the level of validity is relatively low.

---

<sup>111</sup> This is not an example from a PISA test item; it only used to exemplify a fundamental problem. It is told by an American mother. Her daughter, a bright girl who always problematised the multiple choice questions, was stamped as below average.

<sup>112</sup> 16 coders to code science and mathematic and 4 to code reading was required for 4500 and 6000 (OECD 2009: 112). A quick approx calculation gives the need for between 1400 and 2000 coders all together.

### *Presentation*

Presentation is determined by the statistical methods that have been chosen. In PISA a *psychometric model*, which only permits “a one-dimensional variation along the chosen scales” (Dolin 2007: 100) is used. This one-dimensionality is criticised by several of the authors in Hopmann, Brinek & Retzl (2007). Langfeldt (2007: 232) asks “whether the one dimensional scale of a sum score is a valid standard for comparing nations”. Dolin (2007: 113) argues that the “better the evaluation is at capturing complex skills, the more difficult is it to present the results in the form of simple, comparable data”. Sjøberg (2007: 212) states that “PISA (...) is dominated and driven by psychometric concerns, and much less by education”. Bottani/Vrignaud (2005 in Wuttke 2007: 244) argue:

Since the assessment of competences within each of the four subject domains is strictly one-dimensional, any inter-population comparison implies a ranking. This explains the primordial role of league tables in PISA: They are not only a vehicle for gaining media attention, but they are deeply rooted in the conception of the study”

The results of each country are divided into six competence levels according to Gaussian distribution rules, meaning that the percentage of students in the different levels are almost constant (Wuttke 2007: 258). Gaussian distribution makes small differences in mean scores between countries significant, something that might place a country as number thirteen on the league table although only seven countries performed significantly better. This was what happened to Norway in PISA 2000, according to Roe (2006: 179) as referred to in 5.2.3. Taking into account the power of those league tables, this is an important piece of information.

### *Erroneous basis for grounded information?*

“[T]he primary reason for developing and conducting large-scale international assessments is to provide empirically grounded information which inform policy decisions” (OECD 1999: 7). The problems with the test items, sampling, scoring, coding and league tables as discussed in this chapter ought to lead to questions about the trustworthiness of this information. Hence a study where there are uncertainties attached to the reliability and validity also questions the generaliseability, the core stone of a quantitative study. “One cannot generalise test results beyond their area of validity” Dolin (2007: 114) argues. However, the trustworthiness is seldom questioned. From all sorts of researchers, politicians and even some journalists the limitations of the study are underlined; that the league tables

do not tell everything about a country's school system (Hjetland 2009, Kjernsli & Roe 2009). But, as long as the study is used as if the league tables are trustworthy, the reservations do not mean a lot. What if more people start to question how the bricks in the tower of PISA are laid? What if more statistical experts could analyse the fundamental technical issues of PISA and pedagogues and other scholars could analyse the fundamental values, wills and ideas PISA lean on? Taken the power of PISA, it could be wise.

*There are more standardised tests than PISA*

The US has been the hotbed of standardised testing, as argued earlier in this thesis. This is also a country where standardised tests are used extensively (Apple 2006, Au 2009, Klees 2008, Kohn 2000, Sacks 1999). High stake testing has had a skyrocketing development in the US:

Exams used to be administered mostly to decide where to place kids or what kind of help they needed: only recently have scores been published in the newspaper and used as a primary criteria for judging children, teachers, and schools – indeed, as the basis for flunking students or denying them diploma, deciding where money should be spent, and so on. Tests have lately become a mechanism by which public officials can impose their will on schools, and they are doing so with a vengeance” (Kohn 2000: 2).

In other countries the use of standardised tests has not been that extensive, but times are changing. One argument is that international tests and PISA in particular, have been used as a wrecking bar towards more national tests (cf. 2.3.1.). Moreover it is argued that PISA also influences what happens in the classroom more directly. Dolin (2007: 121) argues that “there are signs that PISA, besides exerting an influence on teaching, has also had an influence on the actual objects clause of the elementary schools, so as to direct the teaching to conform to a greater degree with what PISA is capable of measuring”. As many as 71% of the teachers in the survey done by Union of Education Norway agree that many schools probably would use PISA-like tasks to prepare the students for the assessment. Sjøberg (2009) and Niels Christie<sup>113</sup> are worried that PISA might be capable of overriding the Norwegian curriculum plan. Jones *et.al.* (2008: 50) also worry about the capabilities of PISA: “Like any institutionalised system of measurement, PISA did not simply describe facts; it created them and gave them significance. It served to re-regulate schooling by specifying with scientific force a set of objectives for the school, around which reform project could cohere”.

Moreover, PISA is only one of the standardised tests students are exposed to in our time. Some are international; others are regional, national or local. Some are curriculum-based, and some are not. Some are *norm-referenced* and compare the student to all the other students and thus have the aim to sort them in winners and losers, others are *criterion-referenced* and compare students to a set of standards (Kohn 2000). All are summative. In chapter II some of the tests are mentioned, but from the national level only the Norwegian tests are. Actually Norway is not the best example to show how far it might go, although 81 % of the teachers in the survey administrated by Union of Education Norway agreed that the Norwegian school increasingly is characterised by tests.

### 6.3.3 A methodological straitjacket?

Without entering into the discussion of quantitative vs. qualitative research, it seems like the growing number of standardised tests has contributed to one methodology having precedence over the other. Denzin & Lincoln (2005) indicate a growing methodological backlash. With its quantitative approach and huge impact area, PISA can be seen as a part of this backlash, representing a global standard or accountability movement as Saha<sup>114</sup> puts it, as well as a part of the emerging evidence movement, I will add.

#### *The evidence movement*

The focus on evidence based knowledge and practises has led to the establishing of a number of so called evidence producing organisations all over the world, often called *clearinghouses*<sup>115</sup>. Examples of clearinghouses are *Cochrane Collaboration*, the first one in this genre, established in the early 1990s in the field of clinical medicine and *Campbell Collaboration*<sup>116</sup>, established in the late 1990s. The latter has entered among other sectors

---

<sup>113</sup> From a speech in Oslo 22.02.08

<sup>114</sup> Those concepts were used by Lawrence J. Saha from Australian National University in the meeting *Globalisation and comparative Global Pedagogies* 25.03.09 at CIES conference 2009 while characterising PISA .

<sup>115</sup> *Clearinghouse* is a “central institution or agency for the collection, maintenance, and distribution of materials, information, etc” (Utdanningsforbundet 2008: 10).

<sup>116</sup> The head office of Campbell is for the time being located in Oslo:  
[http://www.campbellcollaboration.org/about\\_us/index.php](http://www.campbellcollaboration.org/about_us/index.php) (Accessed 04.12.08)

into the field of education. They do not produce knowledge themselves, but aim to gather the “best” knowledge, to secure the quality of existing research and to make databases and/or systematic reviews and distribute it to professionals within the field (Utdanningsforbundet 2008). Clearinghouses can be seen as important organisers helping researchers or other interested to get hold of the best research from all over the world. They can also be seen as organisations that reinforce the methodological backlash.

While the different clearinghouses operate within a somewhat broader knowledge concept, both the Cochrane and Campbell organisations have an understanding of an evidence hierarchy, where knowledge derived from *randomized controlled trials* (RCT) are regarded more “true” than knowledge gained from other types of research. Approximately 2/3 of the systematized knowledge reviews made by those two organisations relies on research done by RTC methods (Hansen 2007 in Utdanningsforbundet 2008: 16).

#### *What is the “truth”?*

As quoted by Kvale (1996) and Jones (2007) in 6.1.2, there are many reasons why quantitative research is on the offensive. One seductive aspect of such kind of research is all the sophisticated statistical techniques that make it possible to deal with a huge amount of data and generalise the findings to apply to many more cases than those researched. This makes quantitative research powerful. The results of quantitative research are possible to present as numbers that are visualised in graphs, bars and charts. Lay people as well as politicians and journalists are not always capable to understand how the numbers came about; they have to trust the researchers. This makes the researchers powerful. Even not the *outsider* researchers always understand the complicated techniques completely. This makes the *insiders* powerful.

It is, even for experts, rather difficult to understand the statistical and sampling procedures, the rationale and the models that underlie the emergence of even test scores. In practice, one has to take the results at face value and on trust, given that some of our best statisticians are involved. But the advanced statistics certainly reduce the transparency of the study and hinder publicly informed debate (Sjøberg 2007: 212).

It is difficult to argue against a graph – or a league table. The numbers have a clear voice. The trust in the numbers and graphs produced by quantitative methods has developed to be enormous, also in social science. They are seen to represent value-neutral, reproducible, unambiguous and objective facts (Kvale 1996). The thick descriptions delivered by the



qualitative researchers do not have the same force. If read it might be more easy to understand how the results came about, to understand their logic. But they are not able to generalise, and hence not understood as that “true” as the findings delivered by a table.

The notion of what is acceptable knowledge in a field is an epistemological question. Like in ontology we find oppositional poles, one holding a *deductive positivistic* understanding of what is most “true”, another an *inductive interpretivistic* understanding of knowledge as contextual, that subject matters. The knowledge we are constructing from standardised tests like PISA, is built on the notion that objectivistic, positivistic, quantitative knowledge is more “true” than the knowledge built from below.

The strictly formalized procedures of categorization and quantification are ways of ordering and structuring the social world, with quantification as one means of legitimating administrative decisions. In the social sciences, positivism has entailed a philosophic bureaucracy that suppresses the subjective and social dimensions of social research” (Kvale 1996: 64).

## 6.4 Why?

In this chapter I have discussed the issue of setting standards and used the PISA assessment to discuss the questions about whose knowledge and what kind of knowledge that counts. This has been done with a critical eye; both warnings and weaknesses have been found. Naturally there will always be weaknesses connected to a research, however this particular research has had huge impact, something that make the warnings important to take into consideration and the weaknesses important to analyse.

It seems that one impact of PISA is the rise in number of standardised tests as such, and thus a shift toward more emphasis on output. “In the shift from input to output indicators, the PISA surveys were key” Jones *et.al.* (2008: 49) argue. Why this emphasis on output? Output will always vary according to the context; people are unequal by design, school systems differ by choice, cultures are distinct by traditions and countries are different by structures. Why do we have to compare their education output? OECD has given an answer to this question; education should contribute to the production of the “right” human capital in order to raise the economic growth. The comparative tests are supposed to help finding out what works in the production of human capital.

What could the next step possibly be? To standardise education systems all over the world to enable all countries to compete at the same level? Or will standardisation of education in fact lead to an opportunity divide? Like the one we get while inviting an ape, a cow, a goldfish, a tortoise and a bird into a competition where the goal is to reach the top of the tree as quick as possible? Will educational standardisation lead to the production and reproduction of inequalities? American scholars like Apple (2006), Au (2009), Kohn (2000), Kohn and Shannon (2002) and Sacks (1999) who have done research on the impact of the long lasting use of standardised tests in the US, are worried.

Who will be able to become “the most competitive economy in the world”<sup>117</sup> as argued in the Lisbon strategies of EU, if standards set by some sets standards for what counts for the many?

The German economist Friedrich List (in Chang 2002: 4, the italics are original) argued the following in 1841: “It is a very common clever device that when anyone has attained the summit of greatness, he *kicks away the ladder* by which he has climbed up, in order to deprive others of the means of climbing up after him”. The quotation reveals a conscious action. Is what happens in the sector of education also a conscious action? Why do we otherwise create and support systems that place children and also countries in positions where the danger of becoming losers is imminent? Is it due to lack of knowledge, *embedded errors*<sup>118</sup> (Chambers 1997) – or is it done deliberately?

The questions posed on the last pages will not be answered in this thesis, however I want us to have them in mind while turning to the last chapter. There my point of departure will be the emphasis on output and the connection between output and standardised tests

---

<sup>117</sup> [http://europa.eu/scadplus/glossary/lisbon\\_strategy\\_en.htm](http://europa.eu/scadplus/glossary/lisbon_strategy_en.htm) (accessed 01.12.08).

<sup>118</sup> *Embedded errors* sustain mistakes by researchers due to their own professionalism, the distance to the research object and the danger of holding power.

## 7. The bricolage

I have now presented all the bits and pieces that I have gathered from documents, discussions and own experiences in order to answer my research question. What is left is to sew them all together in order to finish the bricolage, something this conclusion intends to contribute to.

My research question asks whether standardised tests might be seen as the pivot in the marketisation of education, if they might be *The Hub in the Wheel*. To answer this question I have tried to see the *marketisation process* and *standardised tests*, the two most central concepts in my question, from different perspectives. I have put them into and tried to analyse them according to historical, ideological, theoretical, economic and political frames.


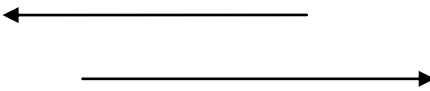
The documents have been breeding and the angle of incidences has been widening through out the process. Moreover, by now I have more questions than answers, insight makes humble. Nevertheless findings have emerged. They might not be consistent enough to give a firm answer to my much too wide research question, but they certainly give some indications in a direction that might confirm my hypothesis as it is formulated in the question.

### 7.1.1 The contradictions between

On the next page I have placed contradictory concepts from the study on each side of a continuum. In dialectics, such concepts are seen as interrelated and even interdependent, although some of them seem mutually excluding. Between the different concept pairs, there are tensions created by different understanding of the concepts and how they relate to reality. These different understandings may cause struggles. Solutions to such struggles are solved in different ways, but the powers involved and the balance between them are decisive when it comes to the outcome. The power balance that determined the development of the economic as well as the education sectors in the years following WWII led to a more human capitalism as well as towards a more progressive and human educational policy in the western part of the world, especially in the 1970s. At that time education was rhetorically seen as a right on its own. Input, material and immaterial, was steadily improved in order to empower as many students as possible. When the power balance between the two superpowers was displaced in the late 1980s and early 1990s, neoliberalism went on the offensive, and capitalism became

rougher. Within education the emphasis changed from input to output alongside with the more outspoken demand for efficiency, accountability and the “right” human capital.

The focus on output made standardised tests the key, and explains the development of a test culture. It is simply impossible to measure output without any measuring tools. However, this does not explain why and how the emphasis has shifted from input towards output and what kind of consequences it has entailed. In the following I will try to conclude by using my findings and the below continuum as a visual tool.

<b>Table 7.1: Contradictory and dialectical concepts</b>	
	
Input	Output
Homo Politicus	Homo Economicus
Cooperation	Competition
Public	Private
Education as a right	Education as a commodity
Multiplicity	One size fits all
Collective	Individual
Interpretivism	Positivism
Constructivism	Objectivism
Qualitative	Quantitative
Inductive	Deductive
Creativity	Adaptility
	

I will argue that the shift in emphasis from input to output was not a choice taken by educationalists or scholars, but by forces outside the education sector. Moreover it is a direct consequence of the guiding ideas held by those forces: Economic growth is seen as the path

leading towards development and prosperity, and the role of education is to contribute to such growth through the production of the “right” human capital. People, Homo Economicus, need competition and control to become proficient and effective, private enterprises are more efficient than public and the invisible hand of the Market is the best way of promoting development. Education is just another commodity that should follow the rule of the Market. Commodities are products and judged as such, hence education should be measured according to what it produces; the output.

To view education as a commodity represents a narrow understanding of knowledge, which is easy to measure and export. Such a position might be in danger to create one-size-fits-all solutions. The rhetoric of the neoliberals and the educational policy that has emerged in the wake of the neoliberal offensive, has had certain epistemological, ontological and methodological implications. They have formed the way knowledge about schooling is produced as well as how knowledge is valued and passed on. The influences and the results might be direct or indirect, intended or unintended, nevertheless it forms the education and what kind of qualifications – and probably what kind of people – that are produced.

It seems that international tests have become important in the search for what works in the production of the “right” human capital and increased efficiency. According to my findings, the national use of international test results has opened up towards more national tests. As long as quasi markets need comparable sizes to function, the national tests can be seen as facilitators to a more business friendly policy in general. Seen from such a perspective the international as well as the national tests, might have contributed to a change in the way people think about education.

### 7.1.2 The connections within

What emerged during my study were the connections between the concepts in each column, the vertical relationship. The neoliberal policy within economic as well as education sector represented by the supranational organisations and the political *right* and gradually also the social democrats, seems to adhere more to the ideas characterised by the concepts in the right column than in the left. The arrows below the continuum are meant to represent development trends caused by the power relations between the *right* and the *left* from the WWII until today. The first one, the shortest pointing toward the left, shows the development during the

1970s. The second, the longer pointing to the right, shows the development trend we have seen the last quarter of a century. The lengths of the arrows indicate that neither earlier nor at the present have the ends of the continuum been reached. We might however be closer to the extreme *right* today than we were to the extreme *left* in the 1970s, thus the neoliberal offensive has been worldwide and has lasted for almost 25 years.

In the change from input to output, standardised tests were keys. I have sketched a figure to make a visual picture of this connection. Since it is only a sketch possible to develop further, I have placed it in appendix J.

- Emphasis on output is directly connected to standardised tests.
- Emphasis on output is indirectly connected to standardised tests through the need for comparison that follows the idea of education as a commodity.
- Emphasis on output is indirectly connected to standardised tests through the idea that Homo Economicus need control.

These three ballpoints make the frame of the four following chains of thought that also are visualised in the sketch.

*The first chain of thought; output - standardised tests:* Without measuring tools, output can simply not be measured. Standardised tests are such tools.

*The second chain of thought; business - output - standardised tests:* Economic growth and to make profit are central capitalistic goals. Education is seen as a way to attain economic growth; however education can also be seen as a way to gain profit. The first argument is often used openly in the debate, the second one is not. Education as a commodity opens up for the possibility to gain profit. My findings have shown that during neoliberalism, private capital has entered into public sector, although in various degrees in different countries. In the case of education, private enterprises have entered into the education sector through private schools, through services offered by private enterprises and through the introduction of ICT. One service is connected to standardised tests. The prize that the public sector or individual parents or students are paying to private enterprises for standardised tests to be formulated, printed, distributed, prepared for, accomplished, coded, scored, analysed and published is unknown. Taken the huge rise in the amount of standardised tests; international,

regional, national and local, the profit most probably is enormous for those involved in business connected to education, also called edubusiness.

*The third chain of thought; business – markets – competition – comparison – standardised tests:* The neoliberal edubusiness require quasi-markets where the commodity can be sold. For such markets to function, competition must be created, and competition needs comparable measures. The results of standardised tests and the publishing of them in particular, the output, might be seen as such measures.

*The fourth chain of thought; business – Homo Economicus – control – standardised tests:* Homo Economicus is playing a key role as a means; regarding education as a commodity, and as a tool; people that need carrots and sticks to “deliver” must be controlled. Standardised tests that have been developed to measure students output are also used to measure performance of teachers, schools and school systems. Students output might thus function as carrots or sticks.

### 7.1.3 *A Hub in the wheel*

I have tried to show that the shift from emphasise on input towards emphasis on output have been important in the neoliberalisation process. In this thesis I have called this process the marketisation of education. In the shift towards emphasis on output, I will argue that standardised tests have been keys directly and indirectly. Standardised tests might thus be seen as an important tool in the marketisation of education.

I have to keep the question mark in the heading of this thesis, however, thus I have realised that a vehicle has more than one wheel. Standardised tests, and particular the *use* of their results, may be seen as *a* hub in one of the wheels moving the neoliberalisation process ahead. This might be important to take into consideration at a time when standardised tests become more and more important, like when the Norwegian *centre/left* government has decided to establish a unit for psychometric research (St.prop. nr.1 2008 – 2009).

Even though not everyone are willing to see the connections between the neoliberalisation process and the use of standardised tests, it is important to remember that choices that are

taken are political and have ideological, epistemological, ontological, methodological as well as practical implications. Those are represented by the contradictory concepts listed in table 7.1. When I started to work as a teacher, the choices that were taken in Norway when education policy was formed were different from today. Nowadays it seems that the argument still is: *There Is No Alternative*. The world has changed, and we just have to adapt to it. However, what has been made by man, might be changed by man. There are always alternatives to development made by man.



## 8. References

- Amin, Samir (2006). The Millennium Development Goals. A Critique from the South. In: *Monthly Review*, vol 57 no.10: 1-15.
- Apple, Michael W. (2006). *Educating the "Right" Way*. 2<sup>nd</sup> edition. New York/London: Routledge.
- Arnove, Robert F. (2003). Reframing comparative education. The dialectic of the global and the local. In: Robert F. Arnove and Carlos A. Torres (eds.). *Comparative education: the dialectic of the global and the local*. Lanham, Md: Rowman & Littlefield: 1-23.
- Asplan Analyse (2005). Omstillinger; Hva koster de? Hva krever de? [Readjustment. What do they cost? What do they demand?]  
<http://www.asplananalyse.no/Analyse/Omstillinger%20-%20Hva%20koster%20de.%20hva%20krever%20de.pdf> (assessed 16.12.08).
- Au, Wayne (2007). Epistemology of the Oppressed: The Dialectics of Paulo Freire's Theory of Knowledge. In: *Journal for Critical Education Policy Studies*, Volume 5, No 2.  
<http://www.jceps.com/index.php?pageID=article&articleID=100> (accessed 16.02.09).
- Au, Wayne (2009). *Unequal by design. High-Stakes testing and the Standardization of Inequality*. New York, London: Routledge.
- Baker, Russ (2002). Stealth TV. In: Kohn, Alfie & Patrick Shannon. *Education, Inc. Turning Learning into a Business*. Revised Edition. Portsmouth, NH: Heinemann: 35- 43.
- Bank, Helene (2006). Utdanning, en tjeneste for hvem? [Education, a service for whom?] In: Birgit Brock-Utne & Liv Bøyesen (eds.). *Å greie seg i utdanningssystemet i nord og sør*. Bergen: Fagbokforlaget: 236-252. [To manage in the education system in north and south.]
- Bergesen, Helge Ole (2006). *Kampen om kunnskapsskolen*. [The fight about the knowledge school]. Oslo: Universitetsforlaget.
- Bond, Patrick (2006). *Looting Africa*. Pietermaritzburg: University of KwaZulu-Natal Press.
- Bredeveien, Jo Moen (2009). Ber Solhjell ligge unna Oslo. [Asks Solhjell to stay away from Oslo] In: *Dagsavisen* 26.02.09: 7
- Breidlid, Anders (2006). Verdensbanken og norsk bistand til utdanning I Afrika – en kritisk kommentar. [The World bank and Norwegian aid to education in Africa – a critical comment.] In: Birgit Brock-Utne & Liv Bøyesen (eds.). *Å greie seg i utdanningssystemet i nord og sør*. Bergen: Fagbokforlaget: 253-269. [To manage in the education system in north and south.]

- Brock-Utne, Birgit (1996). Reliability and validity in qualitative research within education in Africa. In *International Review of Education* 42 (6): 605-621.
- Brock-Utne, Birgit (2006). *Whose education for All? Recolonizing the African mind*. San Diego CA: Africana Books.
- Brock-Utne, Birgit (2007a). Learning through a familiar language versus learning through a foreign language – A look into some secondary school classrooms in Tanzania. In: *International Journal of Educational Development* 27: 487- 498.
- Brock-Utne, Birgit (2007b). Worldbankification of Norwegian development assistance to education. In: *Comparative Education*, Vol 43, No 3: 433- 449.
- Brock-Utne, Birgit (2008a). *Hva var egentlig sosialpedagogikk?* [What was actually social pedagogy?] A speech held on the 70 years anniversary of The Institute of Pedagogical Research at University of Oslo 07.11.08, unpublished.
- Brock-Utne, Birgit (2008b). The effects of the neo-liberal agenda on education in some African countries. In *Research Policy and Praxis*, Volum 6: 91-116.
- Bryman, Alan (2004). *Social Research methods* 2<sup>nd</sup> edition. Oxford: Oxford University Press.
- Chambers, Robert (1997). *Whose reality counts? Putting the last first*. London: Intermediate technology publications.
- Chang, Ha-Joon (2006). *Kicking away the ladder. Development strategy in historical perspective*. London: Anthem Press.
- Chomsky, Noam (2006). *Kicking Away the Ladder: Corporate Economics vs. Democracy*. Institute for Social and Cultural Communicating. A speech held 01.12.2006. Z video production.
- Crossley, Michael & Keith Watson (2003). *Comparative and international research in education: globalization, context and difference*. London: Routledge Falmer.
- Cummings, William K. (2003). *The Institutions of Education*. Oxford: Symposium Books.
- Dale, Roger (2005). Globalisation, knowledge economy and comparative education. In: *Comparative Education*, 41:2, 117-149.
- Denzin, Norman K. & Yvonna S. Lincoln (2005). Introduction. The Discipline and Practice of Qualitative Research. In: Denzin, Norman K. & Yvonna S. Lincoln (eds.). *The sage handbook of Qualitative research*, 3<sup>rd</sup> edition. Thousand Oaks/London/New Dehli: Sage Publications: 1-32.
- Denzin, Norman K. & Yvonna S. Lincoln (2008). Critical Methodologies and indigenous Inquiry. In: Denzin, Norman K. & Yvonna S. Lincoln and Linda Tuhiwai Smith. *Handbook of Critical and Indigenous Methodologies*. Thousand Oaks/London/New Dehli/Singapore: Sage publications: 1-20.

- 
- Dolin, Jens (2007). PISA – An Example of the Use and Misuse of Large-Scale Comparative Tests. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA. Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 93-125.
- Drechsler, Wolfgang (2005). The Rise and Demise of the New Public Management. In *Post-autistic economics review*. Issue no.33: 1-9.  
<http://www.paecon.net/PAERreview/issue33/Drechsler33.htm> (accessed 06.04.08).
- Duffy, Brenden (1987). The Analysis of Documentary Evidence. In: Judith Bell. *Doing your research project*. Philadelphia: Open University press: 53-57.
- Eriksen, Tore Linne (2006). Globalt økonomisk apartheid? [Global economic apartheid?] In: Magnus Marsdal (ed.). *Økonomisk apartheid. Nyliberalismens verdensorden*. ATTAC Norge: 21-63. [Economic apartheid. The world order of neoliberalism].
- Freire, Paulo (1985). *The Politics of Education*. Westport, Connecticut/London: Bergin & Garvey.
- Freire, Paulo (1993). *Pedagogy of the oppressed*. First published in 1970. London: Penguin Books.
- Garbo, Gunnar (2008). Nyliberalismen er verken ny eller liberal. [Neoliberalism is neither new nor liberal.] Kolofon.
- Giddens, Anthony (1998). *The Third Way*. Cambridge/Oxford/Malden MA: Polity Press.
- Giroux, Henry (2003). Selling Out Higher education. In: *Policy Futures in Education*, Vol 1, Number1, 2003: 179-199.
- Gonzales, Eugenio J. & Albert E. Beaton (1994). The Determination of Cut Scores for Standards. In: Tuijnman, Albert C. & T. Neville Postlethwaite (eds.). *Monitoring the standards of education*. Oxford: Pergamon:171-190.
- Gorostiaga, Xabier (1993). New Roles for Universities of the South. Envio: *The monthly Magazine of Analysis on Central America*, Vol.12. No.144: 29-40.
- Harvey, David (2003). *The New Imperialism*. New York: Oxford University Press.
- Harvey, David (2005). *A Brief History of Neoliberalism*. New York: Oxford University Press.
- Hatcher, Richard (2002). *The business of education*. Keynote address at the Annual Course of Association of Education Psychologists, 24 – 26 October, Hinckley, Leicester.
- Hatcher, Richard (2009). *The education policies of the Labour government in England*. A speech held at the welfare conference 2.-3-March 2009, Oslo:  
[http://www.velferdsstaten.no/velferd/makt\\_velferd/Velferdskonferansen.html](http://www.velferdsstaten.no/velferd/makt_velferd/Velferdskonferansen.html)  
 (assessed 23.04.09).

- Hill, Dave (2006). Education Services Liberalization. In: Ellen Rosskam (ed.). *Winners or losers? Liberalizing public services*. Geneva: International Labour Office: 3-54.
- Hjetland, Helga (2004). *GATS og utdanning* [GATS and education]. Introduction to a seminary on GATS and education, Oslo, 11.11.2004:  
<http://www.velferdsstaten.no/attachment/d91bb72b8681562a31bd4957f07efc14/f1504075b95b66253f97a67e01a567b5/Hjetland.pdf> (assessed 24.04.09).
- Hjetland, Helga (2009). Ikkje la PISA dominere opplæringa. [Do not let PISA dominate the training]. In: *Bedre skole* nr.1 2009. Utdanningsforbundet: 29.
- Hofstetter, Rita & Bernard Schneuwly (2008). Bovet's Dilemma – Examinations or No Examinations: the Swiss contribution to the Carnegie initiative. In: Lawn, Martin (ed.) *An Atlantic Crossing?* Oxford: Symposium Books: 83-98.
- Hopmann, Stefan Thomas & Gertrude Brinek (2007). Introduction: PISA According to PISA – Does PISA Keep What It Promises? In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 7-19.
- Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.) (2007). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag.
- Husén, Torsten & Albert Tuijnman (1994): Monitoring Standards in Education: Why and How it Came About. In: Tuijnman, Albert C. & T. Neville Postlethwaite (eds.). *Monitoring the standards of education*. Oxford: Pergamon: 1- 21.
- Hvistendal, Rita & Astrid Roe (2003). *Hvordan gjør elever fra språklige minoriteter det i PISA undersøkelsen?* [How do the students from language minorities manage in the PISA assessment?]  
[http://www.skolenettet.no/moduler/templates/Module\\_Article.aspx?id=24493&epslanguage=NO](http://www.skolenettet.no/moduler/templates/Module_Article.aspx?id=24493&epslanguage=NO) (accessed 28.04.09).
- Hörmann, Bernadette (2007). Disappearing Students. PISA and Students With Disabilities. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit verlag: 157-174.
- Jarning, Harald & Gro Hanne Aas (2008). Between Common Schooling and the Academe: The International Examinations Inquiry in Norway, 1935 – 1961. In: Lawn, Martin (ed.) *An Atlantic Crossing?* Oxford: Symposium Books: 118-202.
- Jensen, Bjarne (2007). *Offentlige velferdsgoder. Økonomistyring, regnskapsprinsipper og New Public Management*. [Public welfare goods. Economic management, accounting principles and New Public Management.] Elverum: Rapport nr.12, Høgskolen i Hedmark.

- 
- Jones, Ken, Chomin Cunchillos, Richard Hatcher, Nico Hirtt, Rosalin Inns, Samuel Josua & Jürgen Klausenitzer (2008). *Schooling in Western Europe*. Hampshire/New York: Palgrave Macmillan.
- Jones, Phillip W. (2007). Education and the world order. In: *Comparative Education* Vol.43, No 3, Routledge Tylor & Francis Group: 325- 337.
- Kincheloe, Joe L. & Peter McLaren (2005). Rethinking critical theory and qualitative research. In: Denzin, Norman K. & Yvonna S. Lincoln (eds.). *The Sage handbook of Qualitative Research*, 3<sup>rd</sup> edition. Thousand Oaks/London/New Dehli: Sage publications: 303-342.
- King, Kenneth (2007). Multilateral agencies in the construction of the global agenda on education. In: *Comparative Education* Vol. 43, No 3, Routledge Tylor & Francis Group: 377-391.
- Kirke - og Undervisningsdepartementet (1987) Mønsterplanen for grunnskolen. [The national curriculum.] Oslo: Kirke- og Undervisningsdepartementet/Aschehoug.
- Ki-Zerbo, Joseph (1994). The Need for Creative Organizational Approaches. In: Diouf, Mamadou and Mahmood Mamdani (eds.). *Academic Freedom in Africa*. Dakar: Codesria Book Series: 26-39.
- Kjærnsli, Marit & Svein Lie (2006). TIMSS og PISA: Prinsipper og hovedfunn fra studiene i 2003. [TIMSS and PISA: Principles and main findings from the studies in 2003.] In: Birgit Brock-Utne & Liv Bøyesen (eds.). *Å greie seg i utdanningssystemet i nord og sør*. Bergen: Fagbokforlaget: 149 – 159. [To manage in the school system in North and South.]
- Kjærnsli, Marit & Astrid Roe (2009). Flere svakheter ved spørreundersøkelse og tolkninger. [Several weaknesses in inquiry and interpretation.] In: *Bedre skole* nr.1. Utdanningsforbundet: 28.
- Klees, Steven J. (2008). A quarter century of neoliberal thinking in education: misleading analysis and failed policies. In: *Globalisation, Societies and Education*, Vol. 6, No 4: 311-348.
- Kohn, Alfi (1992). *No Contest - the case against competition*. Revised Edition. Boston/New York: Houghton Mifflin Company.
- Kohn, Alfi (2000). *The case against standardized testing*. Portsmouth, NH: Heinemann.
- Kohn, Alfie (2002). Introduction: The 500-pound gorilla. In: Kohn, Alfie & Patrick Shannon. *Education, Inc. Turning Learning into a Business*. Revised Edition. Portsmouth, NH: Heinemann: 1-11.
- Kohn, Alfie & Patrick Shannon (2002). *Education, Inc. Turning Learning into a Business*. Revised Edition. Portsmouth, NH: Heinemann.

- Kubow, Patricia K. & Paul R. Fossum (2007). Theory in Comparative Education. In: Kubow, Patricia K. & Paul R. Fossum. *Comparative Education: Exploring issues in international context*. New Jersey: Pearson Merrill Prentice Hall: 31-67.
- Kunnskapsdepartementet (2003). *Nasjonale prøver og gjennomføring våren 2004*. [National tests and implementation, spring 2004.] Rundskriv publisert 16.09.2003 Oslo: Kunnskapsdepartementet.  
<http://www.regjeringen.no/nb/dep/kd/dok/rundskriv/2003/Nasjonale-prover-og-gjennomforing-varen-2004.html?id=109339> (accessed 07.01.09)
- Kunnskapsdepartementet (2006). *Læreplanverket for kunnskapsløftet*. [The National Curriculum for the Knowledge Promotion.] Oslo: kunnskapsdepartementet/Utdanningsdirektoratet
- Kvale, Steinar (1996). *Interviews. An Introduction to Qualitative Research Interviewing*. Thousand Oaks, California: Sage Publications.
- Langfeldt, Gjert (2007). PISA – Undressing the Truth or Dressing Up a Will to Govern? In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit verlag: 225-240.
- Lawn, Martin (2008a). An Atlantic Crossing? The Work of the International Examination Inquiry, its Researchers, Methods and Influence. In: Lawn, Martin (ed.). *An Atlantic Crossing?* Oxford: Symposium Books: 7-37.
- Lawn, Martin (ed.) (2008b). *An Atlantic Crossing?* Oxford: Symposium Books.
- Lawn, Martin (2008c). Blowing up the Citadel of Examinations: the English Committee and the Carnegie Corporation. In Lawn, Martin (ed.). *An Atlantic Crossing?* Oxford: Symposium Books: 39-59.
- Lawn, Martin, Ian Deary & David Bartholomew (2008). Naïve, Expert and Willing Partners: the Scottish Council for Research in Education in the International Examinations Inquiry. In: Lawn, Martin (ed.) *An Atlantic Crossing?* Oxford: Symposium Books: 119-136.
- Lie, Sverre Arne (2009). LO refser SSØ. [LO (The Norwegian Confederation of Trade Unions) reprimand SSØ (The Norwegian Government Agency for Financial Management)] In: *Klassekampen* 04.04.2009: 8.
- Lund, Henrik Herløv (2008 a): New Public Management i Danmark – en prinsipiell kritik. [New Public Management in Denmark – a fundamental critique]. In: *Kritisk Debat*. March. <http://www.kritiskdebat.dk/artikel.php?id=548> (accessed 29.04.08).
- Lund, Henrik Herløv (2008 b): New Public Management i Danmark – en konkret analyse og kritik. [New Public Management in Denmark – a concrete analysis and critique]. In: *Kritisk Debat*. April. <http://www.kritiskdebat.dk/artikel.php?id=551> (accessed 06.04.08).

- 
- Marsdal, Magnus E. & Bendik Wold (2004). *Tredje venstre*. [The third left]. Forlaget Oktober.
- Mazrui, Alamin (1997). The World Bank, the language question and the future of African education. Race & Class. In: *A Journal for Black and Third Liberation*. Vol 38. No3. January – March: 35-49.
- Navarro, Vincent (2004). Inequalities Are Unhealthy. In: *Monthly Review*, Vol. 56 No.2: 26-30.
- Nilsen, Fred Harald (2008). Mer nasjonal styring. [More national steering]. In: *Utdanning* 22: 8.
- Nyerere, Julius K. (1968). Education for Self-Reliance. In: Julius K. Nyerere. *Essay on Socialism*. Dar es Salaam: Oxford University Press: 44-76.
- OECD (1999). *Measuring Student knowledge and Skills. A New Framework for Assessment*. <http://www.oecd.org/dataoecd/45/32/33693997.pdf> (accessed 11.11.08).
- OECD (2007a). *PISA 2006 Science competencies for tomorrow's world*. Vol.2 <http://www.oecd.org/dataoecd/30/18/39703566.pdf> (accessed 04.04.09).
- OECD (2007b). *PISA 2006 Science competencies for tomorrow's world*. Vol.1. <http://www.oecd.org/dataoecd/30/17/39703267.pdf> (accessed 04.04.09).
- OECD (2009). *PISA 2006 Technical Report*. <http://www.oecd.org/dataoecd/0/47/42025182.pdf> (accessed 04.04.09).
- Olsen, Vegard (2007). Large-scale International comparative Achievement Studies in Education: Their Primary Purposes and Beyond. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 265-293.
- Patton, Michael Quinn (2002). *Qualitative Research & Evaluation Methods*. 3<sup>rd</sup> edition. Thousand Oaks/London/New Dehli: Sage Publications.
- Peräkylä, Anssi (2005). Analyzing talk and text. In: Denzin, Norman K. & Yvonna S. Lincoln (eds.). *The Sage handbook of Qualitative Research*, 3<sup>rd</sup> edition. Thousand Oaks/London/New Dehli: Sage publications: 869-886.
- Phillips, Gary W. (1994). Methods and Issues in Setting Performance Standards. In: Tuijnman, Albert C. & T. Neville Postlethwaite (eds.). *Monitoring the standards of education*. Oxford: Pergamon: 191-212.
- Prais, S.J. (2007). England: Poor Survey Response and No Sampling of Teaching Groups. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 139-155.



- Prior, Lindsay (2008). *Using Documents in Social Research*. Los Angeles/London/New Dehli/Singapore: Sage Publications.
- Puchhammer, Marcus (2007). Language-Based Item Analysis – Problems in Intercultural Comparisons. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 127-137.
- Reikvam, Rolf, Bent Sofus Tranøy, Anders Ekeland, Rolf Rønning, May Hansen & Aksel Hagen (2008). *Utveier*. [Expedients.]  
[http://www.unio.no/kunder/unio/mm.nsf/lupGraphics/Utveier\\_031208.doc/\\$file/Utv\\_eier\\_031208.doc](http://www.unio.no/kunder/unio/mm.nsf/lupGraphics/Utveier_031208.doc/$file/Utv_eier_031208.doc) (accessed 16.12.08).
- Reinert, Erik S. (2004). *Global økonomi. Hvordan de rike ble rike og hvorfor fattige blir fattigere*. [Global economy. How the rich became rich and why the poor become poorer]. Oslo: Spartacus.
- Rikowski, Glenn (2007) *Marxist Education Theory Unplugged*.  
<http://www.flowideas.co.uk/print.php?page=243> (accessed 16.02.09).
- Rikowski, Glenn (2008). *Postmodern Dereliction in the Face of Neoliberal Education Policy*.  
<http://www.flowideas.co.uk/?page=articles&sub=Postmodern%20Dereliction%20in%20the%20Face%20of%20Neoliberal%20Education%20Policy> (accessed 16.12.08).
- Roe, Astrid (2006). Hva forteller PISA-undersøkelsen om norske elevers lesekompetanse og lesevaner? [What does the PISA examination tell about the reading competencies of Norwegian students?] In: Birgit Brock-Utne & Liv Bøyese (eds.) *Å greie seg i utdanningssystemet i nord og sør*. Bergen: Fagbokforlaget: 190-203. [To manage in the education system in north and south.]
- Sacks, Peter (1999). *Standardized Minds*. New York: Da Capo Press.
- Samoff, Joel (1996). The Structural Adjustment of Educational Research: Reflections. In: Lene Buchert & Kenneth King (eds.). *Consultancy and Research in International Education*. Bonn: DSE: 65-83.
- Samoff, Joel (2003). Institutionalizing international influence. In: *Comparative education: the dialectic of the global and the local*, Robert F. Arnove and Carlos A. Torres (eds.) Lanham, Md: Rowman & Littlefield: 52-114.
- SAPRIN (2004). *Structural Adjustment: The SAPRIN Report, the Policy Roots of Economic Crisis, Poverty and Inequality*. London: Zed Books.
- Seierstad, Dag (2005). *GATS – tjenesteavtalen til WTO*. [GATS – the WTO agreement on services.]  
<http://www.velferdsstaten.no/attachment/d91bb72b8681562a31bd4957f07efc14/914d3b40ada0e74c4b5520e8934c13ae/Standardforedrag+0305.htm> (accessed 01.03.09).



- Sjøberg, Svein (2006). TIMSS, PISA og norsk skole. [TIMSS, PISA and the Norwegian school.] In: Birgit Brock-Utne & Liv Bøyesen (eds.). *Å greie seg i utdanningssystemet i nord og sør*. Bergen: Fagbokforlaget: 190 – 203. [To manage in the education system in north and south].
- Sjøberg, Svein (2007). PISA and “Real Life Challenges”: Mission Impossible? In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.) (2007). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 203-224.
- Sjøberg, Svein (2009). PISA: bekymringsmelding fra skolen. [PISA: a report of concern from the school.] In: *Bedre skole* nr.1.Utdanningsforbundet: 25-27.
- Small, Robin (2005). *Marx and Education*. Hampshire GU/Burlington, VT: Ashgate.
- Stake, Robert E. (2005). Qualitative case studies. In Denzin, Norman K. & Yvonna S. Lincoln (eds.). *The sage handbook of Qualitative research*, 3<sup>rd</sup> edition. Thousand Oaks/London/New Dehli: Sage Publications: 443-466.
- St. meld.nr. 30(2003 – 2004). *Kultur for læring*. [Culture for learning.] Oslo: Det Kongelige Kunnskapsdepartementet.
- St. meld.nr. 16(2006 – 2007). *...og ingen sto igjen*. [...and no one was left behind.] Oslo: Det Kongelige Kunnskapsdepartementet.
- St. meld.nr. 31(2007 – 2008). *Kvalitet i skolen*. [Quality in school.] Oslo: Det Kongelige Kunnskapsdepartementet.
- St. prop. nr.1 (2004 – 2005).  
<http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stprp/20042005/Stprp-nr-1-2004-2005-/7.html?id=434364> (accessed 07.01.09).
- St.prop. nr.1 (2008 – 2009). <http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stprp/2008-2009/stprp-nr-1-20082009-.html?id=530766> (accessed 21.05.09)
- TNC Gallup (2005). *Undersøkelsen blant rektorer og lærere om gjennomføring av de nasjonale prøvene våren 2005*. [An investigation among head teachers and teachers about the implementation of the national goals the spring of 2005.] Utført på oppdrag av Utdanningsforbundet. Oslo: TNC Gallup.
- UNDP (1990). *Human Development Report 1990*.  
<http://hdr.undp.org/en/reports/global/hdr1990/chapters/> (accessed 11.01.09)
- UNDP (1992). The widening gap in global opportunities. In: *Human Development Report 1994*: 34-47: [http://hdr.undp.org/en/media/hdr\\_1992\\_en\\_chap3.pdf](http://hdr.undp.org/en/media/hdr_1992_en_chap3.pdf) (accessed 10.04.08).
- UNDP (1999). Globalization with a human face. In: *Human Development Report 1999*:1-23: [http://hdr.undp.org/en/media/hdr\\_1999\\_en\\_overview.pdf](http://hdr.undp.org/en/media/hdr_1999_en_overview.pdf) (accessed 17.01.09).

- 
- UNDP (2007/2008). *Human Development Report (2007/2008)*.  
[http://hdr.undp.org/en/media/hdr\\_20072008\\_en\\_complete.pdf](http://hdr.undp.org/en/media/hdr_20072008_en_complete.pdf) (accessed 10.04.08).
- UNESCO (2007). *EFA Global Monitoring Report 2008. Education for All by 2015. Will we make it?* Paris/Oxford: Oxford University Press. Also accessible at  
<http://www.unesco.org/en/education/efareport/> (accessed 30.11.08).
- Uljens, Michael (2007). The Hidden Curriculum of PISA – The Promotion of Neo-Liberal Policy By Educational Assessment. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 295-303.
- Utdanningsdirektoratet (2007). *Rammeverket for Nasjonale prøver*. [The Framework of National tests.] Oslo: Utdanningsdirektoratet.  
[http://www.udir.no/upload/Nasjonale%20prøver/Rammeverk\\_for\\_nasjonale\\_prover\\_2007.pdf](http://www.udir.no/upload/Nasjonale%20prøver/Rammeverk_for_nasjonale_prover_2007.pdf) (accessed 07.01.09).
- Utdanningsforbundet (2008). *Evidens og evidensdebattens betydning for utdanningssystemet*. [The importance of evidence and the evidence debate for the education system.] Temanotat 6.
- Valverde, Gilbert V. (2004). Curriculum Convergence in Chile: The Global and the Local Context of Reforms in curriculum Policy. In: *Comparative Education Review*, 48 (2). 174-201.
- Vulliamy, Graham (1990). The Potential of Qualitative Educational Research Strategies in Developing Countries. In: Graham Vulliamy, Keith Lewin and David Stephens (eds.). *Doing Educational Research in Developing countries: Qualitative strategies*. London: The Falmer Press: 7-25.
- Vuorio-Lehti, Minna & Annukka Jauhianen (2008). Laurin Zilliacus and the “War” against the Finnish Matriculation Examination. In Lawn, Martin (ed.) *An Atlantic Crossing?* Oxford: Symposium Books: 137-156.
- Wahl, Asbjørn (2009). *Velferdsstatens vekst – og fall?* [The rise - and fall of the welfare state?] Oslo: Gyldendal.
- Waldow, Florian (2008). Awkward Knowledge: the German delegation to the International Examinations Inquiry. In: Lawn, Martin (ed.) *An Atlantic Crossing?* Oxford: Symposium Books: 61-82.
- Wuttke, Joachim (2007). Uncertainties and Bias in PISA. In: Hopmann, Stefan Thomas, Gertrude Brinek, Martin Retzl (eds.). *PISA zufolge PISA – PISA According to PISA Hält PISA, was es verspricht? – Does PISA Keep What It Promises?* Wien/Berlin: Lit Verlag: 241-263.
- Yang, Rui (2003). Globalisation and higher education development: A critical analysis. In: *International Review of Education*, 49 (2-3): 269-291.

- 
- Zarrouati, Marc (2008). The Battle of the Baccalauréat: the long forgotten story of a divided committee. In Lawn, Martin (ed.) *An Atlantic Crossing?* Oxford: Symposium Books: 99 - 118.
- Øserk, Kamil (2006). *Opplæringsteori og læreplanforståelse*. [Theory about training and the understanding of curriculum planes.] Revidert utgave. Opplandske Bokforlag.
- Østerud, Svein (2006). PISA- og TIMSS- undersøkelsene. Hvor viktige er de for norsk skole, og hvilke lærdommer kan vi hente? [The PISA- and TIMSS- investigations. How important are they for the Norwegian school?] In: Birgit Brock-Utne & Liv Bøyesen (eds.). *Å greie seg i utdanningssystemet i nord og sør*. Bergen: Fagbokforlaget: 190 – 203. [To manage in the education system in north and south.]

## APPENDIX A

### IEA studies<sup>119</sup>

Years	Name	Abbreviations
1959 – 62	The Feasibility study, also called The Pilot Twelve-Country Study	
1970-71	The Six Subject Study	SISS
1981–83	The Classroom Environment Study	
1985,	The Written Composition Study	
1986 to 2003	Pre Primary Education	PPP
1990-91	The Reading Literacy Study	
1995–96	The Languages in Education Study	LES
1989 and 1992	The Computers in Education Study	COMPED
1999	The Second Information Technology in Education Study - Modul I	SITES
1996–97	The Second Civic Education Study	CIVED
2001-02	The Second Information Technology in Education Study - Modul II	SITES
2002*	The Teacher Education and Development Study	TEDS
2005*	The International Civic and Citizenship Education Study	ICCES
2006	The Second Information Technology in Education Study - Modul III	SITES

\* The start of the data collection

---

<sup>119</sup> Information taken from the official IEA web page: [http://www.iea.nl/brief\\_history\\_of\\_iea.html](http://www.iea.nl/brief_history_of_iea.html) (accessed 04.12.08).

## APPENDIX B

### IEA studies: TIMSS and its precursor<sup>120</sup>

Name or description of assessment study	Participating education systems	Target population	Curricular subject(s) assessed	Year(s) for data collection
<b>FIMS</b> - First International Mathematics Study <sup>121</sup>	12	13 years old + students at pre university	Math	1964
<b>SIMS</b> - Second International Mathematics Study <sup>122</sup>	20	13 years old + final grade of secondary education	Math	1980 - 81
<b>SISS</b> -Second International Science Study <sup>123</sup>	24	10 years old, 14 years old, terminal year in secondary school	Science	1983 - 84
<b>TIMSS 1995</b> - Third International Mathematics and Science Study <sup>124</sup>	45	9 years/grade 3 or 4 13 years/grade 7 or 8 Final secondary education	Math & science	1993-97 - First in a four times cycle
<b>TIMSS-R 1999</b> – Third International Mathematics and Science Study Repete <sup>125</sup>	41	Grade 8 / 13 years	Math & science	1997 - 2001
<b>TIMSS-R</b> – Video study <sup>126</sup>	7	Grade 8	Math & science	1998 - 2001
<b>TIMSS 2007</b> <sup>127</sup>	More than 60	As in TIMSS 2003	Math & science	2007 - 2008
<b>TIMSS ADVANCED 2008</b> <sup>128</sup>	10	Final year of secondary school w/advanced	Math and physics	2007 – 2008

<sup>120</sup> All information in table 2.1 and 2.2 is taken from different pages of the IEA web site accessed 04.12.08: Where there is no footnote, the information is taken from this page: [http://www.iea.nl/brief\\_history\\_of\\_iea.html](http://www.iea.nl/brief_history_of_iea.html) (accessed 04.12.08).

<sup>121</sup> [http://www.iea.nl/fims.html?&no\\_cache=1&sword\\_list\[\]=FIMS](http://www.iea.nl/fims.html?&no_cache=1&sword_list[]=FIMS) (accessed 04.12.08).

<sup>122</sup> <http://www.iea.nl/sims.html> (accessed 04.12.08).

<sup>123</sup> <http://www.iea.nl/siss.html> (accessed 04.12.08).

<sup>124</sup> <http://www.iea.nl/timss1995.html> (accessed 04.12.08).

<sup>125</sup> <http://www.iea.nl/timss1999.html> (accessed 04.12.08).

<sup>126</sup> [http://www.iea.nl/timss-r\\_video.html](http://www.iea.nl/timss-r_video.html) (accessed 04.12.08).

<sup>127</sup> <http://www.iea.nl/timss2007.html> (accessed 04.12.08).

<sup>128</sup> <http://www.iea.nl/timssadvanced20080.html> (accessed 04.12.08).

---

		mathematics courses.		
--	--	----------------------	--	--

## APPENDIX C

### IEA studies: PIRLS

Name or description of assessment study	Participating education systems	Target population	Curricular subject(s) assessed	Year(s) for data collection
<b>PIRLS 2001</b> <sup>129</sup>	35	9 years old/grade 4	Reading literacy achievement and policy and practices related to literacy	2001 – first in a five year cycle
<b>PIRLS 2006</b> <sup>130</sup>	41	As in PIRLS 2001	As in 2001	2005 -06
<b>PIRLS 2011</b> <sup>131</sup>				Planned

<sup>129</sup> <http://www.iea.nl/pirls2001.html> (accessed 04.12.08).

<sup>130</sup> <http://www.iea.nl/pirls2006.html> (accessed 04.12.08).

<sup>131</sup> [http://www.iea.nl/fileadmin/user\\_upload/docs/PIRLS\\_TIMSS\\_promotional.pdf](http://www.iea.nl/fileadmin/user_upload/docs/PIRLS_TIMSS_promotional.pdf) (accessed 04.12.08.)

## APPENDIX D

## OECD studies:

<b>PISA:</b>				
<b>Name or description of assessment study</b>	<b>Participating countries <sup>132</sup></b>	<b>Target population <sup>133</sup></b>	<b>Curricular subject(s) assessed <sup>134</sup></b>	<b>Year(s) for data collection <sup>135</sup></b>
<b>PISA 2000 *</b>	43	15 years old in school	Reading, mathematical and scientific literacy  Main focus: reading literacy	2001
<b>PISA 2003</b>	41	As in 2001	Main focus: mathematical literacy	2003
<b>PISA 2006</b>	57	As in 2003	Main focus: scientific literacy	2006
<b>PISA 2009</b>	67 have signed in	As in 2006		2009
<b>TALIS</b> : Teaching and Learning International Survey programme <sup>136</sup>				
<b>PIAAC</b> : Programme for the International Assessment of Adult Competencies <sup>137</sup>				

<sup>132</sup> All information in table 2.3 come from different pages of the OECD web site accessed 10.12.08:  
[http://www.oecd.org/pages/0,3417,en\\_32252351\\_32235907\\_1\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/pages/0,3417,en_32252351_32235907_1_1_1_1_1_1,00.html) In this table the footnote indicates where the information in the **column** is taken from.

<sup>133</sup> As in previous foot note.

<sup>134</sup> [http://www.oecd.org/document/56/0,3343,en\\_32252351\\_32235907\\_33642296\\_1\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/56/0,3343,en_32252351_32235907_33642296_1_1_1_1_1_1,00.html) (accessed 10.12.08).

<sup>135</sup> As in previous foot note

<sup>136</sup> [http://www.oecd.org/document/0/0,3343,en\\_2649\\_39263231\\_38052160\\_1\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/0/0,3343,en_2649_39263231_38052160_1_1_1_1_1_1,00.html) (accessed 10.12.08).

<sup>137</sup> [http://www.oecd.org/document/35/0,3343,en\\_2649\\_201185\\_40277475\\_1\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/35/0,3343,en_2649_201185_40277475_1_1_1_1_1_1,00.html) (accessed 10.12.08).



## APPENDIX E

Table 2.4. National tests						
	2003/04 <sup>138</sup>	2004/05 <sup>139</sup>	2005 – 2007- No National tests	2007/08 <sup>140</sup>	2008/09	2009/10
Grade 4	- Reading - Maths	- Reading - Writing - Maths - English				
Grade 5				- Reading Norwegian - Maths - Reading English	- Reading Norwegian - Maths - Reading English	- Reading Norwegian - Maths - Reading English
Grade 7		- Reading - Writing - Maths - English				
Grade 8				- Reading Norwegian - Maths - Reading English	- Reading Norwegian - Maths - Reading English	- Reading Norwegian - Maths - Reading English
Grade 9						- Reading Norwegian - Maths - Reading English
Grade 10	- Reading - Maths - English	- Reading - Writing - Maths - English				
Form 1		- Reading - Writing - Maths - English				

<sup>138</sup> According to Bergesen (2006).

<sup>139</sup> According to St.prop.nr. 1 (2004-2005).

<sup>140</sup> This row and the two next is according to the White Paper of Education ( St.meld nr.31 2007/08).

## APPENDIX F

### Early fieldwork questions

Research questions	Back-ground documents	In depth documents	In-formal info	Inter-view
<b>What is the history behind the test culture internationally?</b>				
a. Where and why did they start to use standardised tests?	X			
b. What kind of ideology and political forces promote them?	X			
c. How and why did they disperse?	X	X		
d. Who promote them and who oppose them?	X	X	X	X
e. Impact of the National League Tables	X	X	X	X
<b>What is the history behind the test culture in Norway?</b>				
a. When did we start to join the international test programs?	X			
b. Why did we start?	X	X		
c. What effects did it have?	X	X	X	X
d. Who promote them and who oppose them?	X	X	X	X
e. From National League Tables to National Policy Making	X	X	X	X
<b>What are the effects of standardised tests on:</b>				
a. “what” the schools produce (human capital in a narrow or wide perspective)?	X		X	X
b. “how” the schools produce (methods used in schools)?	X		X	X
c. educational business (do standardization open up for private business in schools)?	X			
d. epistemology (what kind of knowledge become relevant and less relevant)?	X		X	X
e. ontology (whose knowledge counts)?	X			

[illegible]

## APPENDIX H

## In depth documents

In dept documents that I intended to use and/or have used in the analysis of the production and use of standardised tests	Original doc.	"New" doc.	Used as reference
OECD (1999). <i>Measuring Student knowledge and Skills. A New Framework for Assessment.</i>	X		Yes
OECD (2007a) <i>PISA 2006 Science competencies for tomorrow's world. Vol.1.</i>	X		Yes
OECD (2007b). <i>PISA 2006 Science competencies for tomorrow's world. Vol.2.</i>	X		No
OECD (2009). <i>PISA 2006 Technical Report.</i>		X	Yes
Hopmannn, Brinek & Retzl (eds.) (2007). <i>PISA According to PISA.</i>	X		Yes
St. meld.nr. 30(2003 – 2004). <i>Kultur for læring.</i>	X		Not much
St. meld.nr. 16 (2006 – 2007) . . .og ingen sto igjen.	X		Not much
St. meld.nr. 31 (2007 – 2008). <i>Kvalitet i skolen.</i>	X		Not much
St. prop. nr. 1 (2004 – 2004).	X		Not much
Introduction of the Oslo tests.	X		No
Survey done by Union of Education Norway on PISA.		X	Yes
Bergesen, Helge Ole (2006). <i>Kampen om kunnskapsskolen.</i>		X	Yes
Lawn, Martin (2008). <i>An Atlantic Crossing?</i>		X	Yes

## APPENDIX I

### Conferences and seminars

<b>Title</b>	<b>Target group</b>	<b>Where</b>	<b>When</b>
Left forum 2008: Cracks in the Edifice	Intellectuals & activists	New York, US	14-16 March
CIES 2008: Gaining Educational Equity around the World	Academic	New York, US	17- 21 March
IMPLAN 2008: The Implications of Language for Peace and development Conference.	Academic	Oslo, Norway	2-3 May
Evaluating the complex	Academic	Oslo, Norway	29 -30 May
CESE 2008: The World in Europe – Europe in the World	Academic	Athens, Greece	7-10 July
European Social Forum 2008	Intellectuals & activists	Malmö, Sweden	17-21 Sept.
Trade union and social movements – what is in it for us?	Intellectuals & activists	Oslo, Norway	16-17 Oct
Norway Social Forum 2008	Intellectuals & activists	Oslo, Norway	6-9 Nov.
World Social Forum 2009	Intellectuals & activists	Belém, Brasil	27 Jan-1 Feb
Public welfare in the time of NPM	Intellectuals & activists	Oslo, Norway	2-3 March
CIES 2009: The politics of comparison	Academic	Charlston, US	22-26 March

## APPENDIX J

The marketisation of education – a sketch

